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

Title	北陸先端科学技術大学院大学　共有計算サーバ使用成果報告2010
Author(s)	尾崎，泰助；佐藤，幸紀
Citation	Technical memorandum (School of Information Science, Japan Advanced Institute of Science and Technology), IS-TM-2011-001: 1-61
Issue Date	2011-08-02
Type	Others
Text version	publisher
URL	http://hdl.handle.net/10119/9866
Rights	
Description	テクニカルメモランダム（北陸先端科学技術大学院大学情報科学研究科）

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

北陸先端科学技術大学院大学

共有計算サーバ使用成果報告 2010

尾崎泰助, 佐藤幸紀 編

2011 年 8 月 2 日

IS-TM-2011-001

北陸先端科学技術大学院大学

情報科学研究科

〒923-1292 石川県能美市旭台 1-1

要旨

2010 年度に北陸先端科学技術大学院大学において学内で共同利用されている計算サーバや並列計算機を用いて行われた研究の概要および発表論文リストを紹介する.

目次

1. JAIST における共有計算サーバ環境 4

2. 情報科学分野の計算サーバ利用研究 8

鼻腔内流れの解析及び流れの特徴に注目した可視化手法の検討 9

偽腔閉塞型大動脈解離から起因する Ulcer-Like Projection (ULP)の経時変化に関する研究 10

行列行列積を用いた大規模 CFD ソルバの高速化 11

ボクセルベース流体構造連成解析による左心室拍動の解析 13

量子拡散モンテカルロ法を用いた電子状態計算 15

分子結晶多形の第一原理電子状態計算 16

NAT 問題フリーな構造化 P2P ネットワークライブラリ libcage の開発と検証 ... 18

CPU の機能設計段階におけるクロックサイクルレベルシミュレーション 19

非同期型集積システムの低電力テストに向けた基礎研究 21

タイミング調整機構を持つ集積化データパス回路の設計最適化 23

コード実行時における正確なループ階層構造の抽出 25

Routing Performance Evaluation of Hierarchical Interconnection Networks 26

A Prediction-Based Green Scheduler for Datacenters in Clouds 27

A Real-time Sound Field Renderer Based on Digital Huygens' Model 28

Report on Using JAIST's Computational Facilities 29

A Study on Statistical Generation of a Table-of-Contents for Multi-documents . 31

Quantum Monte Carlo calculations of stacking interaction of a DNA base pair . 33

3. マテリアルサイエンス分野の計算サーバ利用研究 34

蛋白質水和水 : 分子動力学計計算による溶媒和自由エネルギー，およびデータマイニング法による水分子の振る舞いの解析 35

生物系のプロトントランスファに関する理論的解析 －酵素触媒機構の解明－ 37
高精度擬ポテンシャル法の開発と表面系への応用 38
第一原理計算プログラム OpenMX の次期バージョン基底関数の検証 40
不均一系 Ziegler-Natta オレフィン重合における活性点のフレキシビリティ 42
オーダーN 大規模密度汎関数法コード:OpenMX の開発と応用 45
一列に穴が開いたグラフェンの電気伝導シミュレーション 47
Adsorption of Carbon-based nanostructures on metallic substrates: a DFT approach to large-scale calculations 49
First-principles study of oxygen reduction reaction and hydrogen oxidation reaction with carbon alloy catalysts 50
Ab-initio study of cyclopentasilane and liquid cyclopentasilane 51
Influence of surface ligands on the electronic structure of Fe-Pt cluster 52
Investigation of two hydrogen terminated graphene doped with Nitrogen 53
The Report on Use of Computing Facilities of JAIST 54

4. 知識科学分野の計算サーバ利用研究 56
Modeling the diversity and log-normality in data 57

5. 2009 年度の計算サーバ利用研究 58
A study on RC4 key collision 59

6. 謝辞 61

1. JAIST における共有計算サーバ環境

情報社会基盤研究センター　佐藤幸紀

北陸先端科学技術大学院大学（JAIST）では，全学で共有利用可能な計算サーバは，その利用者が参加する MPC グループを中心として MPC グループの取りまとめを行う MPC 管理グループと計算機の実務的な運用を担当する情報社会基盤研究センター（以下情報センター）とが親密な連携をとりながら運用されている．情報センターと MPC グループ・MPC 管理グループの関係は参考文献[9]や[10]を参照願いたい．

2010 年度の JAIST における共有計算サーバ環境の更新点及び主だった活動を以下説明する．MPC グループとしては，MPC のホームページ（ http://www.jaist.ac.jp/mpc/ ）から新規ユーザーに MPC メーリングリストへ参加するための登録フォームの運用を 2010 年 7 月より開始した。現在は日本語のみであるが、順次更新していきたいと考えている。

MPC 管理グループは MPC グループのユーザーからの声を吸い上げキュークラスの設定の調整として反映することや mpc メーリングリストにおける利用者間の利用の調停を行っている．2010 年度は 2011 年 3 月 22 日に MPC 管理グループと mpc ヘビーユーザーと情報センターによるミーティングを開催した。情報センターより 2011 年 3 月から稼働開始の 9 ノードの GPU クラスタである pcc-gpu の紹介および利用法の説明を行った後、キューイングシステムの構成について議論を行った。また、情報センターより既設の mpc マシンの稼働率の状況を説明した後、各マシンのキュー構成の見直しや利用方針の確認を行った。議論の結果、PC クラスタの大容量メモリノードにおいて商用 HPC アプリケーションを実行しやすくするための変更を行うことを決定した。また、本ミーティングの最後に情報センターよりペタスケールマシン上で動作するアプリケーション開発の基盤となるライブラリを開発するプロジェクト「Open Petascale Libraries（http://www.openpetascale.org/）」に参加する旨が紹介された。

情報センターは並列計算機の導入，H/W や S/W の運用保守，およびユーザーへの各種サポートを行っている．加えて，並列計算機ユーザーの技術レベルの向上へのサポートの一環として半期に一度程度の利用者講習会を行っている．2010 年度は 6 月に Cray XT5，SGI Altix4700，NEC SX-8，Appro クラスタシステムの利用者講習会を開催した。また、これらの利用者講習会の前に並列機の利用に必要な UNIX の知識を紹介する 30 分間の UNIX 初心者セミナーも開催した。また、11 月には性能解析およびプログラムチューニング・キャッシュ効率化適化をテーマとした Cray XT5 利用者講習会と MPI プログラミングと量子科学計算ソフトウェア GaussianO9 の利用テーマとした Appro クラスタシステムの講習会を開催した． また、情報センターは 2011 年 4 月より情報科学センターから情報社会基盤研究センターに改組されるが、引き続き先端科学技術の研究教育を支援する超並列システムのサービス提供および研究開発と利

用技術の高度化に取り組み、MPC のユーザーにとって快適な計算環境を整備していきたいと考えている。

2011 年 3 月に提供されていた計算サーバの概要を表 1 にまとめた．2011 年 3 月より多倍精度の浮動小数点の大規模 SIMD 演算が可能な GPU を搭載するクラスタである pcc-gpu が導入された。このクラスタシステムは 9 ノードから構成され，ノードあたり 2 基の nVidia Tesla M2050　GPU ボードと AMD の 12CPU コアを持つ CPU が 2 基搭載され、Infiniband 4x QDR によりノード間が接続される。

本報告「北陸先端科学技術大学院大学 共有計算サーバ使用成果報告 2010」は 2010 年度に情報センターから提供されている共有計算サーバを利用した研究の概要とその成果報告である．各ユーザーのニーズを的確に把握し，さらに充実した計算機環境を構築することを目的として，MPC 管理グループと情報センターにより mpc メーリングリストにおいて本報告への協力の依頼を行った．その結果，各著者のご厚意によって情報科学分野から 17 件，材料科学分野から 13 件，知識科学分野から 1 件の報告の提出をいただいた．また、編集の都合で 2009 年度の共有計算サーバ使用成果報告に掲載することができなかった 2009 年度分の利用報告も掲載させていただいた。例年と比べて情報科学分野からの寄稿件数が増加しており、幅広いアプリケーション共有計算サーバが利用されている様子がうかがえる。以上のように，共有計算サーバは基礎的な研究環境の一つとしてますます重要性を増しているといえる．

表 1:JAIST で利用可能な計算サーバ(2011 年 7 月 1 日現在)

機種名	主な仕様
Cray XT5	分散メモリ, スカラー型 CPU: Quad-Core AMD Opteron 2.4GHz (Shanghai) 計算ノード: 　CPU: AMD Opteron 2.4GHz×256×8(19.6TFLOPS) 　メモリ: 16GB×256 ＝ 4TB CPU 間接続: 3D トーラス結合 　帯域幅: CPU-CPU 間　6.4GB/s(HyperTransport) 　CPU-メモリ間　5.3GB/s 　ノードから外部へのデータ転送　7.68×18 ＝ 138.24GB/s(双方向) サービスノード: 　CPU: AMD Opteron 2.4GHz×4×8 　メモリ: 16GB×8 ＝ 128GB

NEC SX シリーズ	共有メモリ型、ベクトル処理 CPU: ベクトル型 102.4GFLOPS/CPU (合計 409.6GFLOPS) メモリ:256GB(共有メモリ) メモリバンド幅:1CPU あたり 256GB/s (合計 1024GB/s) ディスク装置:5TB(RAID6) OS:SUPER-UX(UNIX System V 準拠)
SGI Altix4700	共有メモリ型(ccNUMA) CPU:デュアルコア インテル(R) Itanium2(R) プロセッサ メモリ:24GB×96 台 NUMAlink4(6.4GB/秒)ファブリック結合させた共有メモリ型 合計 96 個のプロセッサ(192 個のコア), 2304GB OS:SUSE Linux Enterprise Server 10 SP1
SGI AltixXE250	分散メモリ型 Master ノード:Intel Xeon 2.8GHz/12MB(8 コア)×1 ノード CPU:Intel Xeon 2.8GHz/12MB×2(8 コア)×4 ノード FPGA:Intel Xeon 2.13GHz/12MB(4 コア)×1 ノード FPGA モジュール:XtremeData Inc XD2260i (Altera Stratix III SE260 FPGA×2)
IBM Cell B.E.	分散メモリ型 CPU: IBM Power5+ 2.1GHz(管理ノード) IBM PowerX 8i Cell 3.2GHz ×2×8 ノード 理論性能:217GFlops×8=1.7TFlops
Apollo PC クラスタ	<Apollo gB222X/1143H> 分散メモリ型 システム全体で 704CPU コア, 2560GB のメモリ (高速演算ノード)64node CPU:Intel Xeon 2.93GHz(Nehalem-EP 4core)×2 メモリ:24GB DDR3 (大容量メモリノード)8node CPU:AMD Istanbul 2.93GHz(Istanbul 6core)×4 メモリ:128GB DDR2 ディスク装置:/work 4.8TB(RAID6 Luster ファイルシステム) OS:Rad Hat 5.4 Infiniband 4×QDR
Apollo GPU クラスタ	分散メモリ型 Appro 1323G2-SM10 全体で 144Core, Memory 288GB 1 ノードの構成: AMD Opteron 6136 (Magny-Cours 8core) * 2 基 nVidia Tesla M2050 * 2 基 32GB DDR3 9 ノードのシステム全体で総理論演算性能 10.6TFlops (CPU 1.3TFlops + GPU 9.3TFlops) Infiniband 4xQDR によるノード間接続 ログインノード(SAS ディスク+NFS ファイルサービス) WORK 領域 2.2TB 主なソフトウェア CentOS 5、PGI Compiler、PBS Professional

参考文献

[1] 佐藤 理史(編),"JAIST における超並列関連研究：1992 年度-1993 年度", 北陸先端科学技術大学院大学　情報科学研究科テクニカルメモランダム,IS-TM-94-0001, (1994).

[2] 佐藤 理史(編),"JAIST における超並列関連研究：1994 年度-1996 年度", 北陸先端科学技術大学院大学　情報科学研究科テクニカルメモランダム,IS-TM-97-3, (1997).

[3] 佐藤 理史(編),"JAIST における超並列関連研究(1997 年度)", 北陸先端科学技術大学院大学情報科学研究科テクニカルメモランダム,IS-TM-98-1, (1998).

[4] 林　亮子(編),"JAIST における並列計算機および計算サーバ利用研究(1998 年度-2000 年度)", 北陸先端科学技術大学院大学　情報科学研究科テクニカルメモランダム,IS-TM-2002-003, (2002).

[5] 林　亮子(編),"JAIST における並列計算機および計算サーバ利用研究(2001 年度)", 北陸先端科学技術大学院大学　情報科学研究科テクニカルメモランダム,IS-TM-2002-004, (2002).

[6] 林　亮子(編),"JAIST における並列計算機および計算サーバ利用研究(2002 年度)", 北陸先端科学技術大学院大学　情報科学研究科テクニカルメモランダム,IS-TM-2003-001, (2003).

[7] 林　亮子(編),"JAIST における並列計算機および計算サーバ利用研究(2003 年度)", 北陸先端科学技術大学院大学　情報科学研究科テクニカルメモランダム,IS-TM-2004-002, (2004).

[8] 林　亮子(編),"JAIST における並列計算機および計算サーバ利用研究(2004 年度)", 北陸先端科学技術大学院大学　情報科学研究科テクニカルメモランダム,IS-TM-2005-001, (2005).

[9] 太田理，尾崎 泰助，佐藤 幸紀(編)，"北陸先端科学技術大学院大学 共有計算サーバ使用成果報告 2007"，北陸先端科学技術大学院大学　情報科学研究科テクニカルメモランダム，IS-TM-2008-002, (2008).

[10] 太田理，尾崎 泰助，佐藤 幸紀(編)，"北陸先端科学技術大学院大学 共有計算サーバ使用成果報告 2008"，北陸先端科学技術大学院大学　情報科学研究科テクニカルメモランダム，IS-TM-2009-001, (2009).

[11] 太田理，尾崎 泰助，佐藤 幸紀(編)，"北陸先端科学技術大学院大学 共有計算サーバ使用成果報告 2009"，北陸先端科学技術大学院大学　情報科学研究科テクニカルメモランダム，IS-TM-2010-001, (2010).

2.　情報科学分野の計算サーバ利用研究

鼻腔内流れの解析及び流れの特徴に注目した可視化手法の検討

情報科学研究科
埴田 翔
利用計算機:　AltixXe, Appro gB222X/1143H.

鼻腔は呼吸における重要な器官であり，３つの甲介で構成される．それ故, 内部構造が複雑であるため, 内部での空気の流れも非常に複雑になっている．鼻腔には気管や肺を保護するために, 吸気の温度を体温近くまで調節する加温冷却機能や乾燥した空気を最適な湿度に調節する機能, 空気中の塵などを取り除く防塵機能, 匂いを感知する嗅覚機能などがある．Keck[1] らは鼻腔内の壁面付近に測定機器を挿入することによって, 鼻腔内部の空気の温度や湿度の測定を行っている．その結果, 吸気される空気が, 鼻腔内の複雑な形状と粘膜の機能によって, 咽頭部に到達するまでに十分に加温加湿されることが明かになっている．

数値流体力学を用いて鼻腔のモデルを構築し, 流れをシミュレーション行うことによってで, 鼻腔内の空気の流れや機能などを解明するための研究が行われている．熊畑 [2] らは鼻腔の壁面における湿度と温度のモデルを構築し, 医療画像から再構築した鼻腔の形状を用いた数値シミュレーションを行っている．熊畑らのモデルでは, 鼻腔の粘膜を模して壁面における温度や湿度のやり取りを再現している．

本研究では, 医療画像から鼻腔の領域の抽出を行った．抽出した鼻腔の領域から３次元形状を再構築し, その形状を用いて鼻腔の計算メッシュを生成した．そして, 流入空気の湿度や温度を変化させて鼻腔内の流れの数値解析を行った．また, 熊畑らの温度と湿度のモデルを用いて鼻腔の数値解析を行い, 副鼻腔内の渦領域を抽出と温度や湿度について検討した結果を国際学会である The 6th World Congresson of Biomechanics[3] にて発表を行った．結果として吸気により鼻孔から流入した空気が鼻腔の複雑な形状により, 十分な温度や湿度に加温加湿されることを明かにした．また, 上顎洞内は渦輪が形成されていることを解明した．

鼻腔内の流れについて λ_2 法を用いて渦領域を抽出し, 渦領域について, Vector や Streamline を重畳して可視化する手法の検討を行い, その成果を可視化情報学会 [4] にて講演発表を行った．鼻腔内の流れでは, 渦などの特徴に基いた詳細な可視化が有効であると考える．

Reference:

[1] T.Keck et al. Humidity and temperature profile in the nasal cavity. Rhinology, vol38,167-171,2000

[2] K.Kumahata et al. Nasal Flow Simulation Using Heat and Humidity Models. Journal of Biomechanical Science and Engineering, vol5,565-577,2010

[3]S.Hanida et al. Examination of Extraction with Vortex Regions in Paranasal Sinus of Human Nose. 6th World Congress of Biomechanics, IFMBE Proceedings Vol 31,Part2,736-739,2010

[4]S.Hanida et al. 鼻腔内流れにおける渦領域の抽出及び可視化, 可視化情報学会 2010 講演文集,241-242,2010

偽腔閉塞型大動脈解離から起因する Ulcer-Like Projection (ULP)の経時変化に関する研究

利用計算機：Altix4700, AltixXE クラスタ，Appro gB222X/1143H

情報科学研究科　松澤研究室　森太志

大動脈解離は，偽腔の血流状態によって偽腔開存型大動脈解離(以下　開存型)と偽腔閉塞型大動脈解離(以下　閉塞型)に分類することができる．それぞれの死亡率について報告されており，閉塞型の方が開存型よりも死亡率が高い傾向にある．死亡原因として，瘤破裂や心タンポナーデなどが原因である．閉塞型の死亡率が高い原因として閉塞型のみにみられる Ulcer-Like Projection (以下 ULP)が考えられる．この ULP が，進展していき瘤化することが報告されており，臨床診断ではこの ULP の経時変化が重要であると考えられている．しかし，ULP の経時変化については分かっていない．経時変化には様々な原因があるが，血液の流れの変化および血流による血管壁への力学的ストレスが関与していると考えられる．我々は，CFD 解析を用いて結果を臨床現場にフィードバックすることで臨床診断の向上に役立つと考えた．

我々は，時系列な医療画像データから形状の再構築をおこない ULP の経時変化について検討をおこなった．その結果，ULP 上部の流れに垂直方向の 2 次流れから流れに水平方向の 2 次流れにおける断面方向を決定し 2 次流れをみたとき，本研究で対象としているすべてのケースにおいて，水平方向の 2 次流れの様相における渦中心の移動は，垂直方向の 2 次流れから決定した方向に一致し，渦中心の軌跡の方向に必ず低い壁ずり応力分布がある傾向がみられた．また，ULP の局所部分で高い圧力分布がみられた．我々の結果は，渦が発生し，壁ずり応力が低いような領域において進展することが報告と一致する．そして，すべてのケースにおいて，2 次流れの渦中心の移動が断面方向と一致することより拡大における予測因子になりうるのではないかと考える．解析症例を増やすことでさらなる検証をおこなう予定である．

最後に，この計算には超並列計算機である SGI 社の AltixXE クラスタ，Altix4700, Appro gB222X/1143H を用いた．これらの計算機上で並列計算をおこなうことで計算負荷の軽減，計算時間の短縮ができた．

研究業績

- 森太志，大竹裕志，眞田順一郎，木村圭一，松井修，渡邊剛，松澤照男．血流解析による時系列データを用いた胸部大動脈解離起因する大動脈瘤への影響．日本機会学会総会 2010 年度年次大会講演論文集（2010). pp.93-94.
- 森太志，大竹裕志，眞田順一郎，木村圭一，松井修，渡邊剛，松澤照男．大動脈瘤の進展予測に関する血流解析．日本流体力学会年会 2010 講演要旨集（2010). pp. 220.

行列行列積を用いた大規模 CFD ソルバの高速化

情報科学科 松澤研究室 西條 晶彦

多数の計算を出来るだけ高速に行わなければならない CFD 計算ではハードウェアの進歩に合わせてプログラムを最適化することが歴史的に行われてきた．現在，HPC 環境の構成がベクトル型 CPU から対称マルチコアプロセッサに移行しており，多くの CFD ソルバにおいてマルチコアを生かした並列化手法が行われている．しかしながら，内積・行列ベクトル積演算で構成された従来のアルゴリズムでは高速キャッシュを持った SMP の能力を活かすことができない．なぜなら，内積・行列ベクトル演算では演算あたりのメモリフェッチが 1 対 1 に対応するため，どれだけ演算規模が大きくなっても低速なメモリ帯域に演算速度が制限されてしまうからである．SMP に使われている高速なキャッシュを活かす場合は行列行列積のような演算密度の高いアルゴリズムを用いる必要がある．

行列行列積は密行列の固有値ソルバにおいて最近用いられており，高速なキャッシュを持つマルチコア CPU マシンにおいて理論ピーク性能の 70% に到達するような非常に高い演算性能を発揮することが知られている [1]．しかしながら，疎行列を扱う CFD の線形ソルバにおいて最も性能を必要とする演算は（疎行列の）行列ベクトル演算や，複雑なメモリアクセスを行う処理であり，行列行列積はほとんど用いられてこなかった．

行列行列積を適用できるような線形ソルバとして，GMRES 法の派生で Z.Bai らによる Newton GMRES 法 [2][3] がある．これは GMRES 法が Krylov 部分空間の基底に逐次的な Arnoldi 過程を用いるのに対して，同時に算出できる Newton 基底を用いるものである．本研究では Newton 基底の算出に行列行列積を用いる．具体的には，ベクトル列の集合の直交化部分において行列行列積を適用する．

ベクトル列の直交化法としては Gram-Schmidt の直交化やハウスホルダー法があり，多くの場合精度と演算量を考慮して，修正 Gram-Schmidt (MGS) が用いられる．一方 MGS に対して，古典 GS (CGS) は演算の順番を組み直すことにより行列行列積が適用出来るため，大規模な問題に対して非常に高い性能を発揮する．本研究では行列行列積を用いた CGS の実装に，横澤らの再帰的古典 Gram-Schmidt (RCGS) 法 [4] を用いた．これは正方行列同士で行列行列積の実行を行うことで高速な直交化を行う．

数値実験

ここでは 2 次元 Poisson 問題に用いられる 5 点近似差分行列を用い，サイズを 2 倍毎に増やして反復数と経過時間をみた．

記憶容量と演算量の問題から Krylov 空間の次元を M において打ち切り許容残差に到達するまで反復を繰り返すリスタート型の Arnoldi GMRES(M) 法と Newton GMRES(M) 法をもちいる．ここでは許容相対残差 10^{-5} とし，$M = 63$ とする．再帰的古典 GS 法のブロックサイズは計算をいくつか試した結果最速であった 1 とした．

計算には Cray XT5 の 1 ノードのみで行い，Cray Scientific Library の Quadcore 版の BLAS ライブラリを用いて 2CPU8 並列で並列実行した．

ベクトルのサイズを 2^{13} から 2^{20} まで変えて Arnoldi GMRES(63) と Newton GMRES(63) を実行した結果，表 1 の結果が得られた．どのサイズにおいても Newton GMRES の反復回数は減少し，実行の総 CPU 時間も短縮されている．図 1 は各サイズ毎の最初の 100 回の反復にかか

表 1　Iteration and CPU time

Vector size	Arnoldi Itr.	time [sec]	Newton Itr.	time [sec]
2^{13}	654	49.3	477	32.0
2^{14}	654	162.1	456	53.7
2^{15}	654	248.6	304	99.8
2^{16}	654	515.3	380	239.0
2^{17}	654	1067.9	310	300.4
2^{18}	654	2359.4	299	536.6
2^{19}	654	6670.9	380	1331.2

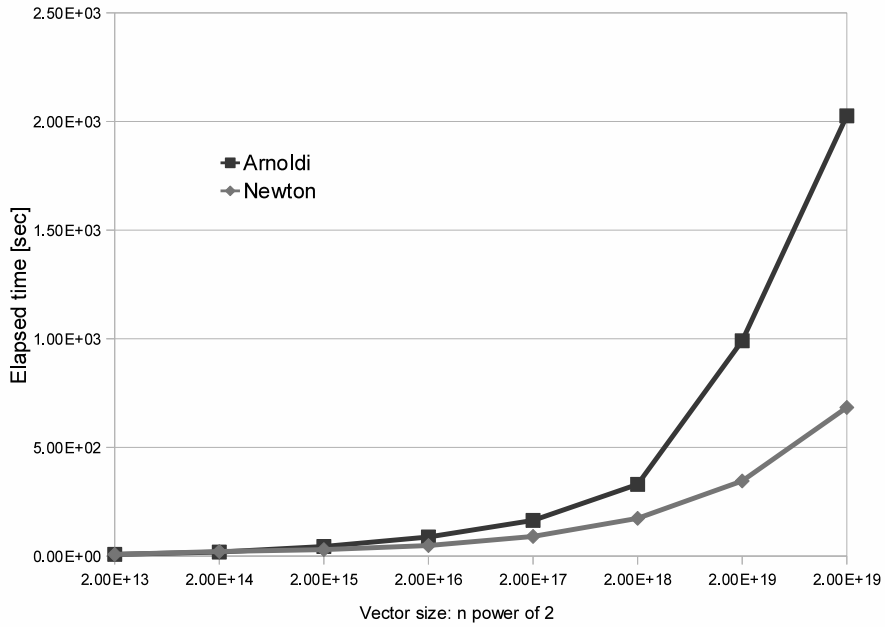


図 1　CPU time for 100 iterations

かった時間を比較している．本研究の手法により行列行列積を使った Newton GMRES では従来の Arnoldi GMRES に比べ最大 3 倍程度の速度向上が達成できた．

[1] 今村俊幸 “T2K スパコンにおける固有値ソルバの開発”, スーパーコンピューティングニュース, Vol.11, No.6,(2009)

[2] Z.Bai, D.Hu, L.Reichel, “A Newton basis GMRES implemenation”, IMA J. Numer. Anal., 14(1994)

[3] J.Erhel, “A parallel GMRES version for general sparse matrices”, Elec. Trans. Numer. Anal., 3(1995),pp160

[4] 横澤, 高橋, 朴, 佐藤, “行列積を用いた古典 Gram-Schmidt 直交化の並列化手法の検討 (数値計算 2)”, 情報処理学会研究報告. [ハイパフォーマンスコンピューティング], 2006

ボクセルベース流体構造連成解析による左心室拍動の解析

情報科学センター 熊畑 清
（現 ㈱富士通長野システムエンジニアリング）

利用計算機: Altix-XE, Apollo PC Cluster

１．緒言

我々は日々の診断や治療を補助するために患者個々人の形状を用いたリアリスティックな血流シミュレーションを，CT データから得られるボクセルデータを直接計算メッシュとして用いて行うことを検討している．血流は心臓の収縮により拍出され，心臓の収縮は心筋細胞の収縮によりなされる．そのためリアリスティックな血流シミュレーションのためには心臓の変形および心筋細胞の振る舞いを考慮した流体・構造・細胞連成解析を行う必要がある．従来流れのシミュレーションは空間に固定された座標系である Eulerian Frame に基づき行われ，固体の変形などのシミュレーションは物質と共に変形する座標系である Lagrangian Frame に基づいて行われてきた．本研究では固体の変形についても Eulerian Frame で流体と同時に解析する手法を開発し，心筋細胞シミュレーションと組み合わせることで左心室モデルの収縮・弛緩による血液の拍出を解析した．

２．手法

本研究で用いている Eulerian Frame で流体運動と固体変形を同時に解く手法では，流体・固体に非圧縮性を仮定し，流体・固体それぞれが 0 から 1 までの比率で混合しているような媒体を想定している．これにより流体・固体それぞれの質量保存式，運動方程式を混合体の質量保存式(1)と運動方程式(2)という一つの方程式系で表すことが可能となる．

$$\nabla \cdot \vec{\boldsymbol{v}}_{mix} = 0 \tag{1}$$

$$\rho_{mix}\left\{\frac{\partial \vec{\boldsymbol{v}}_{mix}}{\partial t} + (\vec{\boldsymbol{v}}_{mix} \cdot \nabla)\vec{\boldsymbol{v}}_{mix}\right\} = -\nabla p + \nabla \cdot \left(\phi_f \boldsymbol{\sigma}'_f + \phi_s \boldsymbol{\sigma}'_s\right) \tag{2}$$

ここで下付添え字の *mix* は混合体，*f* は流体，*s* は固体をそれぞれ意味する．通常固体の応力はひずみ，すなわち変形量と関連付けられる量であるため混合体の速度で表された上記方程式系は閉じず解くことができない．そこで連続体の変形を意味する Cauchy-Green 変形テンソルを，速度勾配テンソルから求める定式化を導出した．これにより応力と速度が結び付けられ方程式系(1)(2)を閉じることができたため流体解析において広く用いられている手法により流体・固体の混合体の運動を解くことが出来るようになる．

式(2)中の固体応力は流体力や外力等による固体の変形により生ずる応力であり，流体力による血管変形などの Passive な変形を扱えるが，心筋組織を構成する心筋細胞が持つ収縮力によるアクティブな変形を扱うためには，収縮力を上記定式化に付与する必要がある．そこで我々は収縮力として心筋細胞シミュレーションモデル”Kyoto モデル”により計算された収縮力を固体の応力項に追加することで心筋組織の自発的な収縮を扱うことに成功した．その際，心筋細胞は配向と呼ばれる心壁の場所毎に異なった方向を持ち，収縮力は配向の向きに作用するため，細胞シミュレーションから得られる収縮力を計算領域全体のグローバルな基底ベクトルと，細胞毎に異なるローカルな座標系の基底ベクトルから定義される回転テンソルを用いて回転することで，心臓の場所毎に配向に応じた収縮力を課すことが可能となった．

３．計算例

上記解析手法を実装し，左心室モデルに対して心壁が収縮力により変形することによる血流駆動のシミュレーションを行った．Fig.1 は左心室を簡易的に模した袋状形状を示す．形状は短軸半径 2.4[cm]，長軸半径 3.6[cm]，厚さ 1.0[cm]の Neo-Hook 体からなる回転楕円体を長軸中央で半分にした形状であり固体は密度 1050.0[Kg/m^3]，ヤング率 50K[Pa]．袋内部と外部は流体とし，内部流体および外部流体はともに密度 1000.0[Kg/m^3]，粘性係数 0.0035[Pa･s]とした．

計算領域の境界条件として回転楕円体を切断する面が接する計算領域上面において内部流体と接する領域のみ Free とし，その他を出入り無の Non-Slip 壁とし，計算領域上面以外の 5 面を Free とした．固体変形の拘束条件は計算領域上面で固体と接する部分は Non-Slip であることから与えられる．収縮力の向きを定める配向ベクトルの初期値として回転楕円体の経度方向を向いたベクトルと，緯度方向を向いたベクトルを基底ベクトルとして，それぞれを 0:100%, 20:80%, 40:60%, 60:40%, 80:20%, 100:0%という比率で混合した 6 つのパターンを用い計算は Altix-XE および Apollo PC クラスターを用い行った．Fig.2(a)-(c)に特徴的な時刻における配向ベクトルの混合比 50%-50%での変形の様子を固体の体積率 0.5 の等値面で示す．収縮力の

増大により袋は収縮し 0.1[s]に付近において最大の変形を示した.以降収縮力の減少に伴い弛緩し 0.4[s]付近においてほぼ元の形状へと戻った事が図中点線で示した補助線により示される.

Fig.3 に配向ベクトルの経度方向の成分と緯度方向の成分の混合比を変えた際の内部流体の駆出率を示す. 経度方向の成分を 0%(緯度方向の成分が 100%)とした際が最も駆出率が高く, かつ実際の心臓に近い結果が得られたが, 実際の心臓の配向に近い 60:40%のケースでは低かった. 実際の心臓では配向は緯度方向を 0 度とした際に, 心臓表面では+80 度, 心臓内側では-80 度となるように, 心壁の厚さ方向に応じた変化をしていることを本モデルではまだ考慮し切れていないためと思われる.

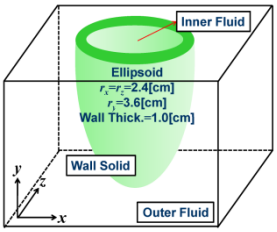


Fig.1 Test Case

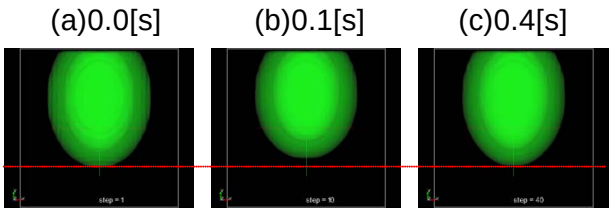


Fig.2 Shape at Typical Time

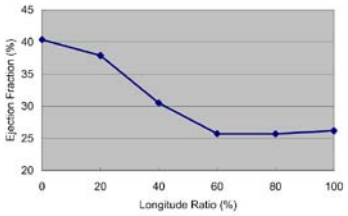


Fig.3 Ejection Fraction

４．むすび

Euler 型流体構造連成解析と，心筋細胞シミュレーションを連成させることで左心室を模す簡易的なモデルにおいて心壁組織の自発的収縮による変形によって流体運動の駆動を再現した. 今後は現コードについての検証を進めつつ, 固体組織の変形の様態が心筋細胞の収縮力へ与える影響を考慮するための開発及び心臓の働きを再現するための検討を行う. ”Kyoto モデル”が心筋細胞内のさまざまな生理学的パラメータを扱えることから，将来的には投薬などによる心筋細胞の振る舞いの変化が，血流に与える影響を解析できるようにすることを期待している. 本研究は理化学研究所「次世代生命体統合シミュレーションソフトウェアの研究開発」プロジェクトからの支援の元行われた. また様々なアドバイスを頂いた「臓器全身スケール研究開発チーム ボクセル班」の諸氏に感謝します.

研究業績

(1) Kiyoshi Kumahata, Shigenobu Okazawa, Akira Amano, Teruo Matsuzawa, "Heart Behavior Simulation on Voxel Based Fluid-Structure Analysis Interacting with Cardiomyocyte Behavior", Proc. The 6th World Congress of Biomechanics in Conjunction with 14th International Conference on Biomedical Engineering (ICBME) and 5th Asian Pacific Conference on Biomechanics (APB), 2008, IFMBE Proceeding Vol.31, pp.461-464, CD-ROM, 00310461, Suntac, Singapore, (2010/08). 査読付き

(2) 熊畑清，西口浩二，岡澤重信，天野晃，松澤照男, "心筋細胞シミュレーターとオイラー型流体構造連成解析による左心室モデルシミュレーション", 日本機械学会 2010 年度年次大会講演論文集, Vol.6, pp.92-92, 愛知県名古屋市(名古屋工業大学), (2010/09).

(3) 熊畑清，岡澤重信，天野晃，松澤照男, "Euler 型流体構造連成解析と心筋シミュレーションの連成による心臓シミュレーション", 日本機械学会計算力学部門第 23 回計算力学講演会講演論文集 CD-ROM- 2102, 北海道北見市(北見工業大学), (2010/09).

(4) 熊畑清，岡澤重信，天野晃，松澤照男, "心臓シミュレーションのための心筋シミュレーションを組み込んだオイラー型流体構造連成解析", 第 21 回バイオフロンティア講演会講演論文集, pp.1-2, 石川県金沢市(IT ビジネスプラザ武蔵), (2010/11).

(5) 熊畑清，岡澤重信，天野晃，松澤照男, "心筋細胞収縮により駆動される左心室拍動のオイラー型流体構造連成解析による解析", 第 24 回数値流体力学シンポジウム講演要旨集, USB, B5-2, 神奈川県横浜市(慶應義塾大学日吉キャンパス), (2010/12).

(6) 熊畑清，岡澤重信，天野晃，松澤照男, "オイラー型流体構造連成解析による左心室拍動シミュレーション", 第 7 回生体工学と流体工学に関するシンポジウム講演論文集, pp.66-71, 石川県金沢市(石川県生涯学習センター), (2011/03).

(7) 熊畑清，岡澤重信，天野晃，松澤照男, "ボクセルベース流体構造連成解析による左心室拍動の解析", 日本機械学会北陸信越支部第 48 期総会・講演会講演論文集, pp.9-10, 長野県上田市(信州大学上田キャンパス), (2011/03).

(8) 熊畑清，岡澤重信，天野晃，松澤照男, "ボクセル型流体構造連成解析による左心室拍動シミュレーション", 理研シンポジウム生体力学シミュレーション研究講演論文集, pp.94-100, 埼玉県和光市(理化学研究所), (2011/03).

(9) 熊畑清, "ボクセルベースの細胞・流体・構造連成で心臓を作りたい", 生命体統合シミュレーションサマースクール, 神奈川県葉山市(湘南国際村センター), (2010/07). ポスター

量子拡散モンテカルロ法を用いた電子状態計算
北陸先端科学技術大学院大学・情報科学研究科　前園グループ

物質デザインにおける複合相関の様相記述は理論的にも旧くから大きな挑戦課題とされてきた。多体電子論の数値的手法である当該法は、この課題に対して最も客観的情報を与える強力なツールであり、大規模高速計算の進歩により、近年、益々、その重要性を増しつつある。本研究では、この手法を用いた非一様系第一原理電子状態計算を用いて、電子物性における高次の電子相関効果の解明を進めている。

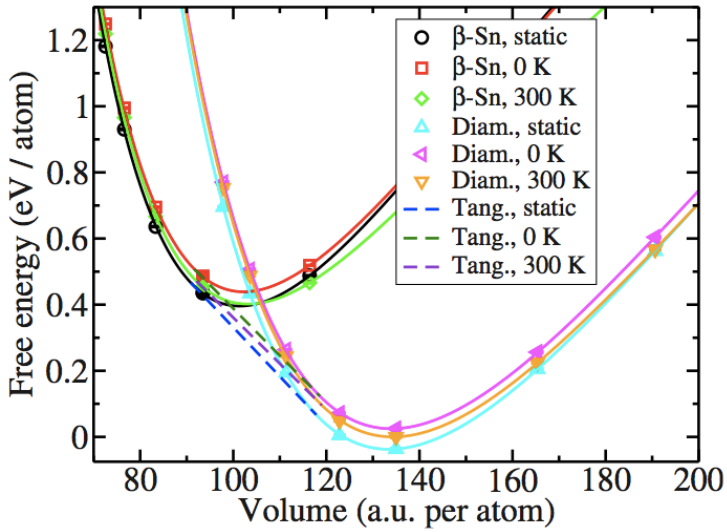


図 4: シリコン固体の構造相転移(ダイアモンド構造→βスズ構造)に関する量子モンテカルロ計算．曲線の極小点から平衡体積が，共通接線から転移を生じる圧力や，その際の体積変化などが算定される [1]．

発表論文：

1) Ryo Maezono, N.D. Drummond, A. Ma, and R.J. Needs, "Diamond to beta-tin phase transition in Si within quantum Monte Carlo", Phys. Rev. B. 82, 184108 (2010).
2) J.R. Trail and Ryo Maezono, "Optimum and efficient sampling for variational quantum Monte Carlo", J. Chem. Phys.133, 174120 (2010).
3) R. Cherian, C. Gerard, P. Mahadevan, N.T. Cuong, and Ryo Maezono, "Size dependence of the bulk modulus of semiconductors nanocrystals", Phys. Rev. B. 82, 235321 (2010).

招待講演：

1) Ryo MAEZONO, "Binding of chromium dimer studied by QMC", PACIFICHEM 2010 (Hawaii, December 15-20, 2010).
2) Ryo MAEZONO, "Electronic structure calculation using Quantum Monte Carlo technique", Ryo MAEZONO, JAIST-CNSI Workshop 2011 (Los Angeles, January 13-14, 2011).

分子結晶多形の第一原理電子状態計算

統計数理研究所　本郷研太
情報科学研究科　前園涼
使用計算機：SX-9

近年，クリーンエネルギーの観点から機能性物質としての分子結晶が注目されており，シリコンに代わる高効率太陽電池材料の候補のひとつとして，注目されている．分子結晶の特徴は，同一分子から構成された結晶でも，それらの構成分子の立体配列の仕方の違い（結晶多形）によって，物理化学的性質が大きく変化することである．この変化は比較的小さいエネルギー（温度換算で 100 ∼ 1000K）で生じる構造相転移に伴う電子状態の変化によって誘起される事が多い．従って，分子結晶の物性を正確に理解して予測するためには，第一原理計算による結晶多形の取り扱いが必要不可欠である．

本研究は，第一原理計算手法として適用範囲の広い密度汎関数理論（DFT）に基づく各種手法によって，分子結晶の電子状態を理解することを目的としている．本研究では，分子結晶のテストケースとして，室温で大きな易動度を示すことが知られている para-diiodobenze (p-DIB) を取り上げる．本研究の DFT 計算は，SX-9 上にインストールされている量子化学計算パッケージ *Gaussian 09* を用いて実行する．Gaussian 09 に代表される量子化学計算パッケージでは，分子軌道（孤立分子系の場合）あるいは結晶軌道（固体周期系の場合）をガウス型関数からなる基底関数セットによって展開する．量子化学計算パッケージが孤立分子系を対象として発展し来たという事情から，利用可能な標準的な基底関数セットは孤立分子を扱うことには適しているが，分子結晶のような固体周期系を取り扱う場合には最適ではあるとは限らない．そこで，本研究では，p-DIB の系に適した基底関数セットを独自に開発するための予備計算を行った．

本研究で扱った基底関数セットは次の 3 つに分類される：(i) 標準的な基底関数セット，(ii) 結晶用の基底関数ライブラリ (http://www.tcm.phy.cam.ac.uk/~mdt26/crystal.html) から入手したセット，(iii) (ii) のグループの基底関数セットを p-DIB 分子結晶の全エネルギーが最小になるように軌道指数を最適化したセット．(i) のグループは，(1) 6-31G (C および H 原子) + 3-21G (I 原子) と (2) 6-311G の 2 つからなり，これらは孤立分子に対して最適化されている．(ii) のグループは，(1) H 原子に (7s1p)/[3s1p]，C 原子に (9s3p1d)/[3s2p1d]，I 原子に (29s20p10d)/[8s7p3d] 縮約関数を用いる基底関数セット（以下ではこれを MDT 基底関数セットとして参照する），および (2) MDT において，I 原子にさらに diffuse 関数として d 軌道関数を加えた基底関数セット（MDT+d）である．ここで，MDT+d セットでは，d 関数の軌道指数のみを全エネルギーに対して最適化している．(iii) のグループは，(ii) の 2 つの基底関数セットを各々最適化したもので，ここではそれぞれ，(1) opt-MDT および (2) opt-MDT+d と記す．(ii) と (iii) の相違点は，ガウス型関数の軌道指数のみであり，基底関数の個数や縮約係数は共通である．各基底関

数セットに含まれるガウス型関数の個数は，6-31G+3-21G セットが６６４個，6-311G セットが 816 個，MDT (opt-MDT) が７８４個，MDT+d (opt-MDT+d) が８２４個となっている．

図 1 に，上記基底関数セットを用いて，DFT の局所密度近似 (LDA) の一つのスキームである SVWN 法により算出した p-DIB 分子結晶の（ユニットセルあたりの）全エネルギーを示す．図 1 に示されるように，孤立分子に対して最適化された標準的な基底関数セットでは，結晶用の基底関数セットと比較して，全エネルギーの記述が非常に悪いことが分かる．特に注目すべき点は，6-311G セットよりも MDT セットの方が基底関数の個数が少ないにもかかわらず，より良い（エネルギーの低い）結果を与えているという点である．すなわち，分子結晶のような固体周期系を量子化学的な計算スキームで扱う場合には，固体周期系に適した基底関数セットを準備する事が重要である事が分かる．次にグループ (ii) と (iii) に着目すると，MDT+d，opt-MDT，opt-MDT+d による結果はほぼ収束しており，これらの変分自由度の枠組みでは最適な基底関数セットが得られたと考えられる．

本研究では，その他の各種 DFT 法に対しても同様の計算を行い，p-DIB 分子結晶に対する基底関数セットについて考察を行う予定である．

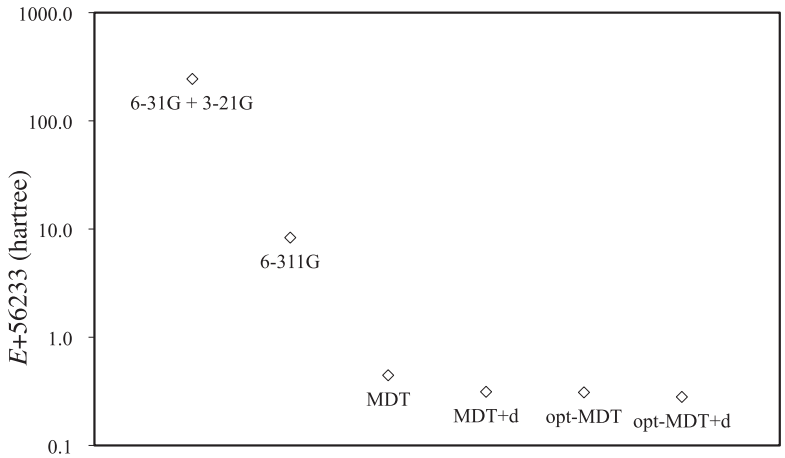


図 1: 各種基底関数セットを用いて得られた SVWN 法による p-DIB 分子結晶の全エネルギー (単位は hartree/cell である)．

研究業績

なし

NAT 問題フリーな
構造化 P2P ネットワークライブラリ libcage の開発と検証

高信頼ネットワークイノベーションセンター
高野 祐輝

構造化 P2P ネットワークはアドレス構造に基づいてメッセージ転送を行うため，Gnutella などの非構造化 P2P ネットワークと比較して，効率的にデータの検索を行うことが出来る手法である．また，規模拡張性にも優れており，参加しているノードが多数となっても効率よくデータ検索が可能である．しかしながら，既存の構造化 P2P ネットワークの設計では，NAT の存在を考慮しておらず，実際にインターネット上で構造化 P2P ネットワークを利用する際には，NAT が原因となる問題に直面する．そこで，本研究では，NAT が存在しても利用可能な構造化 P2P ネットワークの設計を行い，その実証用ライブラリの開発と検証を行った．

インターネット上にはグローバルアドレスを持つノードと，NAT が介在する環境下のグローバルアドレスを持たないノードが存在する．グローバルアドレスを持つノードは，インターネット上からのパッシブなアクセスを受信することが可能であるが，NAT 下に存在するノードはこれが出来ない．そのため，NAT 下に居るノード同士では直接通信が出来なくなり，P2P ネットワークを構成することが困難となってしまう．一般的に，構造化 P2P ネットワークでは全てのノードが全く同じ役割を果すが，本研究ではグローバルアドレスを持つノードと，そうでないノードを分け，階層的にネットワークを構成することで，NAT が介在する環境でも動作可能なように設計した．

NAT を越えるための手法としては，STUN や TURN などが存在するが，これらの手法はクライアント・サーバモデルに基づいて設計されているため，分散環境を前提とする構造化 P2P ネットワークに適用するのは適しているとは言えない．しかしながら，本手法を用いるとサーバに依存せず，分散的に NAT 越えが可能である．そのため，構造化 P2P ネットワークの持つ，規模拡張性や可用性などの利点を保ったまま適用可能であるという特徴を持つ．

本研究では，設計の正しさを実証するために，libcage と呼ぶライブラリの開発を行ったが，その開発と検証の際に，北陸先端科学技術大学院大学が所有する共有計算サーバの一つである Altix 4700 を用いた．libcage は C++ で 18,801 行あり，Boost などのテンプレートライブラリを多くに利用している．そのため，一般的な PC ではコンパイルに数分の時間がかかってしまうが本計算機を用いることで数秒でコンパイル可能になり，開発効率が大幅に上昇した．さらに，構造化 P2P ネットワークは規模拡張性のあるモデルであるが，実装した際には多数のノードでの動作検証が必須となる．通常の PC 環境では数十ノードでの実験が限界であるが，本計算機を用いることで，1 万ノード単位での実験が可能となった．

- 高野 祐輝，井上 朋哉，知念 賢一，篠田 陽一, “NAT 問題フリーな DHT を実現するライブラリ libcage の設計と実装”, 日本ソフトウェア科学会学会誌 『コンピュータソフトウェア』，Vol. 27, No. 4(2010), pp. 58-76
- 高野 祐輝, “インターネット環境に適した構造化 P2P ネットワークソフトウェアの設計と実装”, 北陸先端科学技術大学院大学 博士論文, 2011.3

CPU の機能設計段階におけるクロックサイクルレベルシミュレーション

北陸先端科学技術大学院大学 情報科学研究科 計算機システム・ネットワーク領域 助教

請園 智玲

利用計算機：SGI Altix XE Cluster

１．はじめに

本研究はクロックサイクルレベルシミュレーションを行うことによって，アーキテクチャレベルでの回路の効率性を検証する．クロックサイクルレベルシミュレーションは回路遅延などを考慮せずに考案した回路のクロックサイクル数をカウントするシミュレーション手法である．このシミュレーションは同一のアルゴリズムをいかに少ないクロックサイクル数で完了できるかを LSI 実装作業以前に検証し，実際に物理設計する価値のある回路であるか否かを判断するために用いる．このクロックサイクルレベルシミュレーションは LSI 設計工程で行うシミュレーションの中で最も 1 回のシミュレーションに要する計算量が低い工程であるが，実験的な回路の価値を全方位から調べるために，シミュレーションパラメータを変更して繰り返し実行する回数が非常に多い．このことから，結果的に最終的な回路の価値を判断するための計算量が膨大なものとなる．

２．計算対象 CPU アーキテクチャ

本研究では主に 2 種類の CPU をクロックサイクルレベルシミュレーションによって性能を評価した．1 つ目はバッテリ駆動型組込みシステムの稼働時間向上を狙った低消費電力プロセッサである．本アーキテクチャは今日の CPU の大半を占める SRAM 回路のリーク電力削減を行うために，パワーゲイティング制御を動的に行う．このアーキテクチャは CPU 上で実行するプログラムに依存してその効果が変動するため，多種のプログラムコードをシミュレータの入力として変化させ，多数のシミュレーションを行う必要がある．また，SRAM を用いる記憶回路の構成を変化させた場合も同様にその効果が変動するため，複数の記憶回路構成でシミュレーションする必要がある．2 つ目は低資源プロセッサにおける性能劣化緩和を狙ったプロセッサである．本アーキテクチャは組込み向け CPU の資源制約でキャッシュメモリが十分に確保できない場合のアーキテクチャである．通常，キャッシュメモリの容量に対し性能向上率は比例しない．キャッシュメモリは過度な容量を投入しても一定以上の高性能化が見込めない特性と，容量が少しでも不足する場合に急激に性能を下げる特性をもつ．本プロセッサは後者のキャッシュメモリ容量が不足する場合の性能低下を緩やかにし，キャッシュメモリ投入分の性能の確保を目的としている．

３．評価

2010 年度は２節で示したプロセッサを考案し，そのプロセッサのクロックサイクルレベルシミュレーションを行った．シミュレーションは SPEC 2000 ベンチマークセットと

MiBench マークセットで提供されるアプリケーションをワークロードとして用いた．これらのベンチマークはそれぞれ数十アプリケーションを含み，計測はそれぞれアプリケーション毎に複数の提案ハードウェアのセッティングと性能比較対象ハードウェアのシミュレーションを行う必要がある．この計算量は膨大なものとなるため，高性能並列計算機である SGI Altix XE Cluster を使用し，1 プロセスで 1 セッティングのシミュレーションを行うジョブを多量に発行する計測手法を本研究は採用した．このシミュレーションで，より細かいセッティングの変化を行った場合，どういう特性で性能が変化するかを知ることが可能となった．このシミュレーションから，低消費電力プロセッサでは，提案機構を導入することで，最大で 94%，平均で 52%の L2 キャッシュリーク電力削減効果を得ることができるという結果を得た．一方，性能劣化緩和プロセッサでは，提案機構を導入することで，最大で 1.65 倍，平均で 1.21 倍の性能向上を得ることができるという結果を得た．

４．まとめ

特に計算機アーキテクチャの研究分野では，上述の機能レベルのシミュレーションを繰り返し，プログラム実行時のハードウェアの解析とそれに基づく発見的アイデアを組込んだアーキテクチャの提案が数多くなされている．このようなシミュレーションは時間方向にも空間方向にも分割できず，1 タスク内で十分な並列度を得ることが難しい．このため，1 タスクの完了に非常に長い実行時間を要する．しかしながら，入力パラメータを変えた多数のシミュレーションタスクの実行が必要とされる作業である特性上，Altix XE Cluster のような粗粒度の大規模並列計算機が有効な計算資源となる．今後も情報社会基盤研究センターの大規模並列計算機を利用して計算機アーキテクチャの研究を進めていきたい．

研究業績

１．組込みプロセッサ向けデータキャッシュ制御方式の検討
請園　智玲，田中　清史
計算機アーキテクチャ研究会報告 2010-ARC-191(1),page 1--7, 2010.

１．組込みプロセッサ向け命令キャッシュ制御方式の検討
請園　智玲，田中　清史
組込みシステムシンポジウム 2010 論文集，page 81--86,2010.

２．Reduction of Leakage Energy in Low Level Caches
請園　智玲，田中　清史
Proceedings of The Workshop on Low Power System on Chip (Soc), onlined by IEEE, 2010.

３．動的最適化機構の電力最適化への適用
請園　智玲，田中　清史
先進的計算基盤システムシンポジウム 2010 論文集，page 259--266,2010.

非同期型集積システムの低電力テストに向けた基礎研究

情報科学研究科 助教　岩垣 剛

使用計算機：SGI Altix XE250

研究の背景と目的

本研究で対象としている非同期式集積回路は，現在の主流である同期式の集積回路で問題となっているクロックスキュー問題や消費電力問題等を根本から解決するものとして注目を集めている．今後，非同期式集積回路を用いたシステムがますます広まると半導体技術ロードマップで予想されており，高信頼なシステムの実現には，そのテスト技術が欠かせない．現状のテスト技術が抱える大きな課題の一つは，「集積回路でテスト時に消費される電力をいかに抑えるか」である． テスト時の過度な消費電力は，回路に重大なダメージを与えるだけでなく，誤った（良品を不良品とみなしたり，逆に，不良品を良品とみなしたりするような）テスト結果を生じさせるため，この問題の解決がシステムの信頼性を保証する上で鍵になる．本研究では，非同期式集積回路に対しては議論されていなかった低電力テストについて考察し，その基礎理論の構築を目指した．

研究成果

現在までにさまざまな非同期集積システムの実現方式が提案されているため，各方式に応じた議論が必要となる．本研究では，その糸口として CHAIN とよばれる有力な非同期相互結合網の一方式を取り上げ，考察をおこなった．本年度（2010 年度）は，消費電力を考慮したテスト手法の開発の着手までには至らなかったが，研究を進める中で明らかとなった既存手法の問題点を改善し，今後の議論の土台となる以下の成果を得た．

CHAIN のテストでは， CHAIN を構成する基本モジュール内の故障に対するテストパターンを予め生成しておき，それを CHAIN の外部から対象モジュールに供給する必要がある．このとき，「各テストパターンをいつ，どういう順番で，どの経路を用いて供給するか」という，ある種のスケジューリング問題が生じる．本研究では，テスト時にのみ利用可能な CHAIN の特別なデータ転送方法に着目し，従来は実現できなかったようなテストスケジュールを生成する方法を考案した．具体的には，CHAIN を構成するパイプラインラッチ，完了検出部，経路制御部のそれぞれに対して，先の特別なデータ転送を考慮したスケジューリング問題を定式化し，コーン（CHAIN 内の相互に関係する素子の集合）という概念を用いたテストスケジューリング法を提案した．ベンチマーク回路を用いた計算機実験によって，提案手法が従来手法と同等のテスト品質を保ちつつ，最大で約 40% のテスト時間を削減できることが明らかになった．この傾向は CHAIN の構造が複雑になるほど顕著になると予想している．

本研究では，さらに，上記のスケジューリング問題を異なった観点から解くことを試みた．具体的には，先に説明した三つのスケジューリング問題を整数計画問題へモデル化することをおこなっ

た．整数計画問題へモデル化することで，近年，その性能の向上が目覚しい整数計画ソフトウェアを用いて最適なテストスケジュールを得ることができる．この整数計画モデルに基づく手法によって，テスト時間がさらに改善されることを実験的に明らかにし，先に提案した手法で得られたテストスケジュールがほぼ最適なものであることも確認した．

さまざまな規模や構造の CHAIN に応じて，上記の二つの提案手法をうまく組合せたり使い分けたりすることで，柔軟に効率的なテストスケジュールを決定することができる．さらに，整数計画モデルに基づく手法は，今後，消費電力を考慮したテストを実現するための基礎となる．つまり，整数計画モデルがもつ柔軟性を利用して，テスト時の消費電力に関する新たな制約式や目的関数を既存のモデルに導入することによって，消費電力を考慮した CHAIN のテストスケジューリングが可能となる．

集積回路のテストでは，チップの信頼性を保ちつつ，テストに要するコストを下げることが求められる．本研究の成果は，テストコストの重要な一要素であるテスト時間の削減に寄与するものであり，低コストで信頼性の高い集積回路の実現につながる．

研究業績

1. Tsuyoshi Iwagaki, Eiri Takeda and Mineo Kaneko, “Test scheduling algorithms for delay-insensitive chip area interconnects based on cone partitioning,” *Proc. 3rd International Workshop on Impact of Low-Power Design on Test and Reliability (LPonTR ’10)*, 2 pages, May 2010.
2. Tsuyoshi Iwagaki, Eiri Takeda and Mineo Kaneko, “An approach to test scheduling for asynchronous on-chip interconnects using integer programming,” *Proc. 11th IEEE Workshop on RTL and High Level Testing (WRTLT ’10)*, pp. 69-74, Dec. 2010.

タイミング調整機構を持つ集積化データパス回路の設計最適化

情報科学研究科 金子峰雄
使用計算機 altix-xe

数十ナノメートルの最小線幅で回路構造が形成される現代の極微細集積回路においては、製造ばらつきによる集積回路の特性ばらつきが顕在化しており、タイミング誤りによる製造歩留まり低下やマージン設計による性能向上の飽和が問題となっている。この問題を回路システムと設計の立場から根本的に解決する手法として、製造後にチップ個別のタイミング調整を行う回路方式を提案している。この回路方式が実際のアプリケーションに対して十分な性能を発揮するためには、タイミング調整を予め考慮した集積回路の設計最適化が必要となる。

特に本研究では、アプリケーションの計算実行を行うデータパス部の設計を対象とし、計算結果を保持するレジスタや信号の流れを制御するマルチプレクサへの制御タイミングのクロックタイミングからのずれ(スキュー)を製造後調整する方式を検討した。始めに、この検討対象の方式において回路の性能歩留まりが、回路構造と遅延情報、アプリケーションの回路へのマッピング(資源割り当て)情報から生成される辺重み付有向グラフ(スキュー制約グラフ)が正サイクルを持たない最小クロック周期の確率分布として与えられることを明らかにし、この方式の下での設計最適化目標をスキュー制約グラフが正サイクルを持つ確率(正サイクル確率)の最小化に置くことの重要性と妥当性を指摘した。しかし一方で、この正サイクル確率の厳密な計算の困難さも明らかになり、厳密な評価に代わって近似的に正サイクル確率を大きくする危険なサイクルを最小化する設計問題を考案し、これを数学的な定式化としての有向辺/無向辺混合グラフの順序彩色問題に帰着した。次に、この順序彩色問題のＮＰ困難性に鑑み、同問題を整数線形計画問題(ＩＬＰ)に帰着させて解くことを試み、ＩＬＰ定式化の導出とＩＬＰソルバを用いた解生成実験を行った。

高位合成の代表的ベンチマークである Jaumann Wave Filter (1 周期内 17 演算、ループパイプラインスケジュール：繰返し周期 10 ステップ)と 5th-Order Elliptic Filter (1 周期内 34 演算、ループパイプラインスケジュール：繰返し周期 16 ステップ)について合成実験の概要を紹介する。先ず始めにそれぞれの計算アルゴリズムについてのパイプラインスケジュールを求め、次いで提案手法による資源割り当てのためのＩＬＰ式を生成する。ここで、前者についてのＩＬＰ式は 583 個の変数が出現する 10,921 本の制約式となり、後者については 1,473 の変数、63,911 本の制約式であった。これらＩＬＰ式を代表的なＩＬＰソルバである IBM 社 CPLEX version 12.1 にて解を求め(Jaumann Filter の求解に 10 分から 20 分程度を要し、Elliptic Filter は求解を 1 時間で打ち切った暫定解)、資源割り当てを得ている。こうして得られたデータパス回路合成結果に対して、遅延特性ばらつきに対する性能歩留まりをモンテカルロシミュレーションにより算出した結果を図１、２に示す。これは所定のクロック周期(横軸)において製造後スキュー調整が成功する確率(縦軸)を表しており、従来設計に対して、本手法による設計が、性能歩留まりを大幅に改善している様子がわかる。

提案手法は、本来評価すべきスキュー制約グラフにおける正サイクル確率を正しく評価・最適化できておらず、回路の最適性を保証できていない。今後、より正サイクル確率に近くかつ計算可能な設計最適化目標の設定とそれに基づく回路最適化に取組む予定である。またこの過程で、より効率的にＩＬＰ解を生成できるＩＬＰ定式化も併せて検討する予定である。

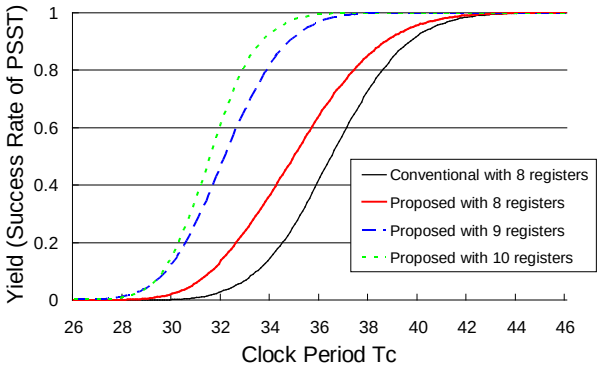


図 1. Jaumann Wave Filter データパス回路のスキュー調整成功確率(歩留まり)

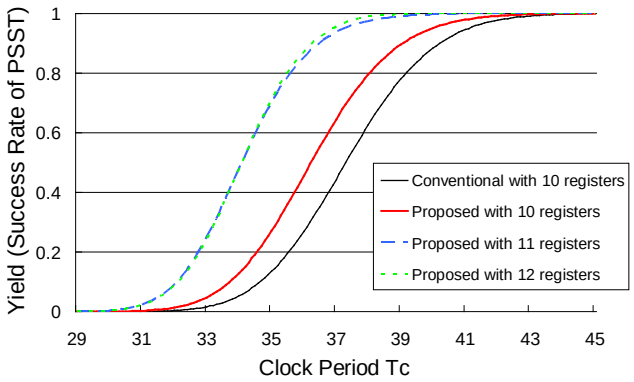


図 2. 5th Order Elliptic Filter データパス回路のスキュー調整成功確率(歩留まり)

研究発表

[1] Mineo Kaneko, ``Ordered Coloring for Skew Adjustability-Aware Resource Binding,'' 電子情報通信学会 ＶＬＳＩ設計技術研究会, VLD2010-42, pp.1-6, 2010.

[2] Mineo Kaneko, ``ILP Approach to Extended Ordered Coloring for Skew Adjustability-Aware Resource Binding,'' 電子情報通信学会ＶＬＳＩ設計技術研究会, VLD2010-75, pp.131-136, 2011.

[3] Mineo Kaneko, ``A Complete Framework of Simultaneous Functional Unit and Register Binding with Skew Scheduling'', Proceedings of the 12th International Symposium on Quality Electronic Design, IEEE Catalog No. CFP11250-CDR, ISBN:978-1-61284-912-6, pp.189-195, 2011.

[4] Mineo Kaneko, Keisuke Inoue, ``Ordered Coloring-Based Resource Binding for Datapaths with Improved Skew-Adjustability'', Proceedings of the 2011 Great Lakes Symposium on VLSI, ACM Order No. 477118, ISBN:978-1-4503-0667-6, pp.307-312, 2011.

コード実行時における正確なループ階層構造の抽出

情報社会基盤研究センター　佐藤幸紀

使用計算機 pcc, pcc-m, alitx-xe

プログラムのソースコード解析やコンパイラ技術に基づく静的フロー解析の枠を超えてコントロールフローを抽出し効率的にハンドリングすることを目的として、動的バイナリ変換（Dynamic Binary Translation）システムにて実行時に正確なループ構造をバイナリコードより抽出する機構を構築した。ループ階層構造部の設計の過程でループネスト構造を正確に抽出するためには reducible loop に分類される自然ループ(natural loop)だけを想定するのだけではなく、irreducible loop の検出も可能なループ抽出法に基づく静的解析によりループ構造を抽出するアルゴリズムを用いて実行バイナリコードを実行する際にループの動的な挙動をモニタリング可能な機構を設計した。

実装したシステムの評価の結果、ソースコード上のループ構造と矛盾しない正確な構造が抽出されていること、バイナリコード実行中随時ループネスト構造が追跡されること、関数呼び出しをまたぎ動的に実行されるループネスト構造を L-CCT（Loop-Call Context Tree)として効率的に表現可能であること、SPEC CPU ベンチマークにおいて実行時ループ検出なしのネーティブのコード実行と比べて平均 4 倍程度のオーバーヘッドにて実行時ループの検出が可能であることが示された。

加えて、各ループのイテレーションの数(ループとリップカウント)を実行時に検出する手法を設計し、SPMD（Single Program Multiple Data）の並列モデルで記述される MPI プログラムにおいて並列化されているループ部分をコード実行時に検出することを試みた。評価実験の結果、ループにおいて並列化されている MPI プログラムから並列部分をコンパイル時のオプション指定やソースコードの参照をすることなく並列化されているループ部分を検出できることを確認した。

研究業績等

[1] Yukinori Sato, Yasushi Inoguchi and Tadao Nakamura. On-the-fly Detection of Precise Loop Nests across Procedures on a Dynamic Binary Translation System. Proceedings of 2011 ACM International Conference on Computing Frontiers. 査読有. 2011.

[2] 佐藤幸紀, 井口寧, 中村維男. 動的バイナリトランスレーションによるループネスト検出とプログラムチューニング支援への応用. 第 18 回ハイパフォーマンスコンピューティングとアーキテクチャの評価に関する北海道ワークショップ（HOKKE-18）. 2010 年 12 月 16 日（北海道）

Routing Performance Evaluation of Hierarchical Interconnection Networks

School of Information Science, Inoguchi Lab., JAIST
Used Machine: Altix 4700
M.M. Hafizur Rahman

Abstract:
The dynamic communication performance (message latency and network throughput) of a massively parallel computer systems consisting of thousands of nodes is evaluated using a routing algorithm. Routing algorithm coordinates the data transfer among the nodes in an interconnection network. The efficiency of a routing algorithm in a network is characterized by low latency and high throughput. Evaluation of the efficiency of a routing algorithm in a hierarchical interconnection network using a personal computer is possible; however, it takes infeasible long time to evaluate it. High speed computation environment of JAIST significantly reduced the simulation time in our experiment. We used our simulator written in C code and run it on Altix 4700 using gcc compiler. Some of the recent results are published in the following journals and conference proceedings.

Publications:

① M.M. Hafizur Rahman, Yukinori Sato, and Yasushi Inoguchi, "On Nonuniform Traffic Pattern of Modified Hierarchical 3D-Torus Network" IEICE Transactions on Information and Systems, Japan, vol.E94-D, No.5, pp.1109-1112, May 2011.

② M.M. Hafizur Rahman, Yukinori Sato, Yasuyuki Miura, and Yasushi Inoguchi, "Dynamic Communication Performance of Hierarchical 3D-Torus Network" Proc. of the 10th PDCN, IASTED, pp. 9 – 16, Innsbruck, Austria.

③ M.M. Hafizur Rahman, Yukinori Sato, and Yasushi Inoguchi, "High Performance Hierarchical Torus Network" Proc. of the 13th ICCIT, pp. 210-215, Dhaka, Bangladesh.

④ M.M. Hafizur Rahman, Yukinori Sato, and Yasushi Inoguchi, "Dynamic Communication Performance Enhancement in Hierarchical Torus Network by Selection Algorithm" Proc. of the 13th ICCIT, pp. 210-215, Dhaka, Bangladesh.

⑤ M.M. Hafizur Rahman, Yukinori Sato, Yasushi Inoguchi, "Dynamic Communication Performance of a Modified Hierarchical 3D-Torus Network under Nonuniform Traffic Patterns" Proc. of the 1st ICNC, IEEE, pp. 167-172, Hiroshima, Japan.

A Prediction-Based Green Scheduler for Datacenters in Clouds

Truong Vinh Truong Duy (0820006)
Inoguchi Lab, School of Information Science
Machine in use: Altix-XE

With energy shortages and global climate change leading our concerns these days, energy consumption of datacenters has become a key issue. Obviously, a substantial reduction in energy consumption can be made by powering down servers when they are not in use. This work aims at designing, implementing and evaluating a Green Scheduler for reducing the energy consumption of datacenters in Cloud computing platforms.

The Green Scheduler is composed of four algorithms: prediction, ON/OFF, task scheduling, and evaluation algorithms. The prediction algorithm employs a neural predictor to predict future load demand based on historical demand. According to the prediction, the ON/OFF algorithm dynamically adjusts server allocations to minimize the number of servers running, thus minimizing the energy use at the points of consumption to benefit all other levels. The task scheduling algorithm is responsible for directing request traffic away from powered-down servers and toward active servers. The performance is monitored by the evaluation algorithm to balance the system's adaptability against stability.

For evaluation, we perform simulations with two load traces. The simulations were conducted on SGI Altix XE nodes during the period from January 2010 to September 2010. The results show that the prediction mode with a combination of dynamic training and dynamic provisioning of 20% additional servers can reduce energy consumption by 49.8% with a drop rate of 0.02% on one load trace, and a drop rate of 0.16% with an energy consumption reduction of 55.4% on the other. Our method is also proven to have a distinct advantage over its counterparts. More experiments are being carried out on Altix XE for the purpose of improving our work.

[1] T.V.T. Duy, Y. Sato, and Y. Inoguchi, "A Prediction-Based Green Scheduler for Datacenters in Clouds", IEICE Transaction on Information and Systems, Vol.E94-D, No.9, 11 pages, Sep. 2011

[2] T.V.T. Duy, Y. Sato, and Y. Inoguchi, "Improving Accuracy of Host Load Predictions on Computational Grids by Artificial Neural Networks", International Journal of Parallel, Emergent and Distributed Systems, Taylor & Francis, 16 pages. To appear in print; published online on 09 August 2010

A Real-time Sound Field Renderer Based on Digital Huygens' Model

Tan Yiyu
Inoguchi Lab, School of Information Science
Used Machine: Pcc-m1

Modeling acoustical behavior in a room is complicated and computationally intense. Although many methods have been proposed for computer simulation and computers have become faster and faster, the procedure is deeply time-consuming as sound space increases. With the development of semiconductor technology, an alternative solution is to simulate and rebuild the sound field distribution using hardware. Compared with other methods, the method based on digital Huygens' model (DHM) is simple and easily implemented by hardware. In this work, two schemes of DHM are derived to analyze sound propagation in a two-dimensional space, and are implemented by field programmable gate array (FPGA). In a two-dimensional space of size 2m × 2m, surrounded by rigid walls, the hardware simulation results of both schemes agree well with the software simulation result, and the calculation errors are relatively small. Compared with the software-based FDTD, original scheme, and updated scheme, a hardware system with 1024 nodes implemented by a Xilinx FPGA chip XC5VLX330T-FF1738 speeds up 7.1 times, 5.5 times, and 3.6 times, respectively.

In this work, synthesis and implementation of the hardware system are carried out in the environment of server pcc-m1 since it has large memory and many processor cores. The time taken by system verification and implementation is shortened significantly by using this server.

Publications

1. Tan Yiyu, Yasushi Inoguch, Eiko Sugawara, Makoto Otani, Yukio Iwaya, Yukinori Sato, Hiroshi Matsuoka, Takao Tsuchiya, "A Real-time Sound Field Renderer Based on Digital Huygens' Model", Journal of Sound and Vibration (in press).
2. Tan Yiyu, Yasushi Inoguch, Eiko Sugawara, Makoto Otani, Yukio Iwaya, Yukinori Sato, Hiroshi Matsuoka, Takao Tsuchiya, "A FPGA Implementation of the Two-Dimensional Digital Huygens' Model", 2010 International Conference on Field Programmable Technology (FPT), pp. 304 – 307.

Name: Ngo Bach Xuan

Student ID: s1120004

Email: bachnx@jaist.ac.jp

School: Information Science

Laboratory: Natural Language Processing (Shimazu's Lab)

Report on Using JAIST's Computational Facilities

Used Machines: Altix4700, Cray XT5

1. Recognition of Requisite Part and Effectuation Part in Law Sentences

Abstract

Analyzing the logical structure of a sentence is important for understanding natural language. In this research, we present a task of Recognition of Requisite Part and Effectuation Part in Law Sentences, or RRE task for short, which is studied in research on Legal Engineering. The goal of this task is to recognize the structure of a law sentence. We investigate the RRE task regarding both the linguistic features and problem modeling aspects. We also propose solutions and present experimental results in a Japanese legal text domain. We got 88.58% with a supervised learning model and 88.84% with a semi-supervised learning model in the F_1 score on the Japanese National Pension Law corpus.

Publications:

[1]**Ngo Xuan Bach**, Nguyen Le Minh, Akira Shimazu. RRE Task: The Task of Recognition of Requisite Part and Effectuation Part in Law Sentences. To appear in *International Journal of Computer Processing Of Languages* (IJCPOL), 2011.

[2]**Ngo Xuan Bach**, Nguyen Le Minh, Akira Shimazu. Exploring Contributions of Words to Recognition of Requisite Part and Effectuation Part in Law Sentences. In *Proceedings of the 4th International Workshop on Juris-Informatics* (JURISIN), pp. 121-132, Tokyo, Japan, November 2010.

[3]**Ngo Xuan Bach**, Nguyen Le Minh, Akira Shimazu. Recognition of Requisite Part and Effectuation Part in Law Sentences. In *Proceedings of the 23rd International Conference on the Computer Processing of Oriental Languages* (ICCPOL), pp. 29-34, San Francisco California USA, July 2010.

[4]**Ngo Xuan Bach**, Nguyen Le Minh, Akira Shimazu. Recognition of Requisite Part and Effectuation Part in Law Sentences. In *Proceedings of 16th Annual Meeting of Association for Natural Language Processing* (NLP), pp. 35-38, Tokyo, Japan, March 2010.

2. **A Semi-Supervised Learning Method for Vietnamese Part of Speech Tagging**

Abstract

This paper presents a semi-supervised learning method for Vietnamese part of speech tagging. We take into account two powerful tagging models including Conditional Random Fields (CRFs) and the Guided Online-Learning models (GLs) as base learning models. We then propose a semi-supervised learning model for both CRFs and GLs for tagging models. The main idea is to use of a word-cluster model as an associate source for enrich the feature space of discriminate learning models for both training and decoding processes. Experimental results on Vietnamese TreeBank data (VTB) showed that the proposed method is effective. Our best model achieved the accuracy of 92.60% and 92.20% when testing on VTB and an independent test, respectively.

Publications:

[1] Le Minh Nguyen, **Xuan Bach Ngo**, Viet Cuong Nguyen, Quang Nhat Minh Pham and Akira Shimazu. A Semi-Supervised Learning Method for Vietnamese Part of Speech Tagging. In *Proceedings of the 2nd International Conference on Knowledge and Systems Engineering* (KSE), pp. 141-146, Hanoi, October 2010.

In these works, we used Cray XT5 and Altix4700 machines to train Brown word clusters and Conditional random fields models.

A Study on Statistical Generation of a Table-of-Contents for Multi-documents

NGUYEN Viet Cuong

Natural Language Processing Laboratory (Shimazu-Shirai Lab)
School of Information Science, JAIST
cuongnv@jaist.ac.jp

June 3, 2011

1 Resources

- Machines: Cray XT3, Cray XT5, SGI Altix 4700, Appro gB222X/1143H.
- Queue Class: SMALL.
- Tasks:
  - Perform word-clustering on the Wikipedia English (about 1.2 billion words).
  - Perform topic-modeling on the Wikipedia English.
  - Generating a table-of-contents for multiple documents.

2 Outline of Research

Generating a table-of-contents (TOC) for multiple documents written about the same topic is a new task in natural language processing. It could play as a navigation tool to help readers quickly locate interesting parts. In addition, readers could look through a TOC to get an overview of the topic of the document set. In our research, we propose a framework for generating a TOC for multi-documents which involves three sequential tasks:

- **Text Segmentation** is a task of splitting a document into topically coherent segments. All documents in the set are put into a text segmentation system to get a collection of segments.
- **Segment Combination** is a task of merging and combining all the segments to form a hierarchical structure of segments (a tree of segments) which reflects the hierarchical structure of information.
- **Title Generation** is a task of generating a title for each node in the tree of segments. A title is a phrase which reflects the content of segments belonging to the node.

In our research, we aim to improve the performance of the above three tasks by using supportive knowledge in terms of semantic and topical information. The supportive knowledge which is a kind of semantic knowledge has been acquired from a large collection of plain texts by unsupervised learning algorithms such as word clustering and topic modeling.

The major research problems and our contributions are summarized as follows.

- First, the task of generating a TOC for multi-documents is new. Therefore, we proposed a framework which integrates the above three tasks in a pipeline to receive a set of documents as the input and to produce a TOC as the output. This framework allows us to improve the performance of tasks individually.
- Second, we focus on improving the performance of the unsupervised linear text segmentation. The current works on the task are mainly based on the assumption of lexical cohesion which consists of reiteration and collocation relations. However, those works only take into account the first type

of relations which can be easily recognized by observing the repetition of words. The second type of relations consists of the systematic and non-systematic semantic relations, which are the most complex relations to be recognized. In our research, we investigate on linguistics phenomena to find that supportive knowledge could be used to recognize these relations effectively. In addition, we also generalized current unsupervised text segmentation methods in a unique framework. The evaluation on public corpora shows the advantages of our model over the current state-of-the-art models.

- Third, the current learning models for the title generation task are still using non-semantic features about words in a text such as frequency, position, part-of-speech, syntactic function, and so on. That may be the reason of the low quality of generated titles of current models. In our research, we investigate on a method to integrate semantic and topical information to the title generation learning model by using supportive knowledge. In addition, due to the lack of training data, we also investigate on using the word clustering to avoid the sparseness of data. We evaluated our proposed approach on a public dataset and get potential results.

- Finally, we investigate on the segment combination task which is raised from our framework for TOC generation for multi-documents. In our research, we proposed a combination algorithm which is based on the hierarchical agglomerative clustering (HAC) method. This algorithm combines segments by the topical similarity between segments. The output of the algorithm is a tree which reflects the hierarchical structure of information. We also proposed a heuristic algorithm to flatten the binary tree which is the output of the HAC-based algorithm to make the output look more realistic.

In summary, main contributions of our research are to propose a framework for generating a TOC for multi-documents and to investigate on using supportive knowledge to improve the performance of the text segmentation and title generation tasks. The improved systems have been evaluated on the public datasets in comparison to the current state-of-the-art methods. We also did experiments on real datasets and employed human to evaluated results.

3 Publications

1. V. C. Nguyen, L.M. Nguyen, and A. Shimazu. Learning to Generate a Table-of-Contents with Supportive Knowledge. *IEICE Transactions on Information and Systems (Special Section on Knowledge Discovery, Data Mining and Creativity Support System)*, Volume 94-D, Number 3, pages 1–9, March 2011.

2. V. C. Nguyen, L. M. Nguyen, and A. Shimazu. Improving Text Segmentation with Non-systematic Semantic Relation. *Computational Linguistics and Intelligent Text Processing (CICLing 2011)*, Lecture Notes in Computer Science (LNCS), Volume 6608–6609, Springer-Verlag Berlin Heidelberg, February 2011. To appear.

3. L. M. Nguyen, X. B. Ngo, V. C. Nguyen, Q. N. M. Pham, and A. Shimazu. A Semi-supervised Learning Method for Vietnamese Part-of-Speech Tagging. In *Proceedings of the Second International Conference on Knowledge and Systems Engineering (KSE 2010)*, pages 141–146, Ha Noi, Viet Nam, October 2010.

4. V. C. Nguyen, L. M. Nguyen, and A. Shimazu. A Semi-supervised Approach for Generating a Table-of-Contents. In *Proceedings of the 7th International Conference on Recent Advances in Natural Language Processing (RANLP 2009)*, pages 313–318, Borovets, Bulgaria, September 2009.

5. V. C. Nguyen, L. M. Nguyen, and A. Shimazu. A Semi-supervised Model for Table-of-Contents Generation. In *Proceedings of the 11th Conference of the Pacific Association for Computational Linguistics (PACLING 2009)*, pages 147–152, Sapporo, Japan, September 2009.

The report on use of JAIST's computational facilities

Nguyen Thanh Cuong
(Collaboration with Prof. Ryo Maezono, School of information science)
Used Machine: SX9, Altix-4700

Quantum Monte Carlo calculations of stacking interaction of a DNA base pair

Weak noncovalent interactions such as van der Waals (vdW) dispersion forces and hydrogen bonds play a central role in biomacromolecules like DNA, RNA, and proteins. However, the accurate description of these interactions remains a major challenge for *ab initio* electronic structure theories. Recently, many methods such as vdW-density functional theory, random phase approximation within the adiabatic connection fluctuation dissipation theorem have been developed to overcome this challenge. Here we studied the stacking interaction of a large biomacromolecule: Adenine-Thymine (A-T) base pair by using quantum Monte Carlo (QMC) method that gives high accuracy results for small systems. We also estimated stacking energy of A-T base pair by using other advanced methods: vdw-density functional theory, Møller-Plesset perturbation theory. We find that QMC method give the reliable results, and well agreement with experiment and vdw-density functional theory

For this work, we use density functional theory and Møller-Plesset perturbation theory calculations with Gaussian software on SX9 machine. The quantum Monte Carlo calculations are carried out on Altix-4700 with 16 cpus/job.

Publications

N. T. Cuong, and R. Maezono, *J. Chem. Phys.* In preparation

3.　マテリアルサイエンス分野の計算サーバ利用研究

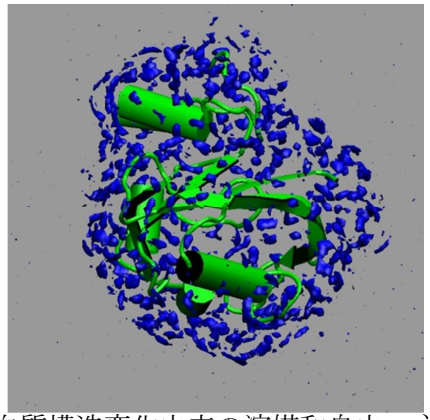
2010 年度　北陸先端大共有サーバー使用報告書　マテリアル科学研究科　助教　水上　卓

「蛋白質水和水：分子動力学計計算による溶媒和自由エネルギー，およびデータマイニング法による水分子の振る舞いの解析」

使用サーバー：Cray XT5，SGI Altix4700

蛋白質は生理的条件下で，3 次元立体構造を保持し，外部刺激に影響されて構造変化し，機能発現の為の化学反応を起こす．一連の現象において，非常に大きな役割を果たしているのが，水分子である．溶媒として蛋白質分子の周辺に存在している水分子は水和水と呼ばれ，これまでに実験，理論，計算的手法のそれぞれにおいて，物理化学的に特異な性質が報告されてきた．そこで蛋白質の機能発現過程における水分子の寄与の解明を目的に，分子動力学計算からのトラジェクトリデータを用いて，1)溶媒和自由エネルギー計算，2)溶媒水分子ダイナミクスのデータマイニング法による分類，という二つのアプローチで研究を進めた．

PYP 周辺の水分子密度分布（1g/mL の 5 倍密度を表示）



1) エネルギー表示法による溶媒和自由エネルギー計算

溶媒和自由エネルギー計算には，エネルギー表示(ER)法を採用した．Amber10 を用いて水溶液中の蛋白質の分子動力学計算を行い，溶質(蛋白質分子)と溶媒(水分子)との相互作用から見積もりを行った．蛋白質は Photoactive Yellow Protein (PYP)を用い，N 端領域と PAS ドメインからなる open-close 構造に着目し，その変化過程におけるエネルギーの変化を見積もった．その結果，open 構造において溶媒和自由エネルギーの数十 kcal/mol 程度の増大を見た．これは常温溶媒中・平衡状態で，10ps 程度の時間間隔において観測できる蛋白質構造変化由来の溶媒和自由エネルギーと同程度であった．我々は，水和水の解析を行い，第一水和水の濃度の変化と溶媒和自由エネルギーが強い相関を持つことを見いだした．

2) 溶媒水分子ダイナミクスのデータマイニング法による分類

溶液中の水分子の運動は非常に揺らぎが大きいことが知られている．それは分子同士の静電相互作用が長距離力であり，かつ相互作用のエネルギーが大きいために，水分子の配位変化において広範囲の水分子が関与するエネルギーのやりとりを伴うからである．このような性質をもつ分子集団に対し，動力学的にその性質を予測したり分類したりすることは，原理的に困難である．我々は，分子動力学データに対して，サポートベクターマシン(SVM)， 主成分解析(PCA)，階層的クラスタリングなどのデータマイニング法を適用し，一旦物理モデルに捕らわれない形で水分子の振る舞いを特徴空間上で取り扱い，データ駆動型の解析を行うことを提案してきた．現在までに，水分子と蛋白質との距離情報を用いずに，個々の水分子の拡散情報のみから水和水とバルク水との分類を試み，一定の成功を納めた．本研究では，水分子の水素結合の配位数の情報を元にして水分子の分類を行い，同様に水和水とバルク水を分けることに成功した．今回得られた情報は，分子の拡散とは別の基底に属する情報であると考えており，今後複数の情報を組み合わせることにより特徴空間の多次元化を図り，より高効率・高精度な解析を目指す．

【国内学会】

1) 分子科学討論会 (2010/9/15)

Photoactive Yellow Protein の溶媒和自由エネルギー計算

水上 卓, 齋藤 大明, 長尾 秀実

2)生物物理学会年会(2010/9/20)

Free Energy Analysis of Photoactive Yellow Protein

Taku Mizukami, Hiroaki Saito, and Hidemi Nagao

【国際学会】

1) Pacifichem 2010 (Honolulu, USA, 2010/12/19)

Taku Mizukami, Ayumu Sugiyama, Dam Hieu Chi, Ho Tu Bao

Title: Classification of protein hydration water by means of data mining method

2) The 52th Sanibel Symposium (Florida, USA, 2011/2/28)

Solvation effect on the structural change of a globular protein: a molecular dynamics study

Taku Mizukami, Hiroaki Saito, Shuhei Kawamoto, Takeshi Miyakawa,

Masashi Iwayama, Masako Takasu, and Hidemi Nagao

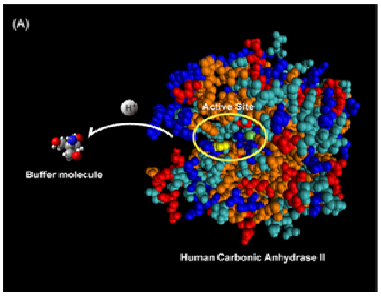
【発表論文】

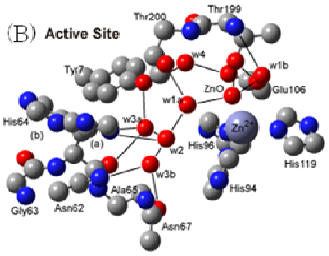
1. Solvation effect on the structural change of a globular protein: a molecular dynamics study Taku Mizukami, Hiroaki Saito, Shuhei Kawamoto, Takeshi Miyakawa, Masashi Iwayama, Masako Takasu, and Hidemi Nagao (submitted) (査読有)

生物系のプロトントランスファに関する理論的解析 －酵素触媒機構の解明－

ナノマテリアルテクノロジーセンター　助教　島原秀登

目的：本研究は,炭酸脱水酵素(CA)をはじめバクテリオロドプシンやシトクローム c オキシダーゼ等自らの機能発現に必須な生物系水仲介型プロトントランスファ(PT)過程について,機能発現に係る水素結合相互作用様式を解明することでその基本原理を探索する構想をもち,具体的に,亜鉛含有 CA 触媒反応の金属イオン触媒と His 残基による一般酸･塩基触媒の観点から,その側鎖イミダゾールの互変異性平衡,酸-塩基平衡,及び水素結合能の NMR 情報を反応速度論解析と相関させた提唱機構について,溶液･結晶構造情報に基づく量子化学計算を行うことにより実験-理論間の整合性を検証し,取得最適化構造に含まれる情報から触媒機構を完全に記述することを目的とする.





概要：CA は二酸化炭素と水を H^+ と重炭酸イオンに変換する亜鉛含有酵素である.その機能発現に枢要なはたらきを担う His64 について,これまでその配向変化を必要としない PT 過程を含む触媒機構の妥当性を評価する計算を行ってきた.そこでは(A)に示すように酵素の活性部位から外部のバッファ分子へ H^+ が移動する内容が含まれる.その(A) から,生化学的な実験によって重要とされる亜鉛とその周辺残基,水分子の酸素を抽出したものが(B)であり,その原子間距離から尤もらしいと考えられる水素結合ネットワークを実線で示した.今年度,特に,Gaussian09 実装 SX-9 マシンによる量子化学的手法(DFT,SMD 法)を用いて,この結晶構造活性部位モデルに付加した水素の位置について構造最適化計算の結果を解析した.その結果,Tyr7 の OH 基が w3a 酸素として示した水と水素結合する可能性が高いことが示され,その OH 基の存在によって His64-$(H_2O)_n$-Zn^{2+}領域の水素結合が安定化されていることが予想された.Tyr7 を Phe に変異させること(Y7F 変異体)によってその OH 基を除いた場合,その変異体の酵素活性は野生体のそれと比較して上昇することが反応速度論解析により知られる.その場合において His64-$(H_2O)_n$-Zn^{2+}領域の水素結合に如何なる変化が生じるのか,報告者は NMR 法を用いた His 残基互変異性/水素結合解析手法によってこれをさらに追跡する実験を行った.その結果,Y7F 変異体の His64 は互変異性定数(K_T)と酸塩基解離定数(pKa)はそれぞれ 0.7 と 6.8-6.9 を示すとの実験データを得た.その K_T 値について,低分子イミダゾール誘導体解析が示す K_T 値と水素結合状態の間の 2 つの関係「i)K_T=0.4 付近を示す His の Nδ1 窒素は水素結合を形成しない状態, ii)K_T=0.6 以上の Nδ1 窒素は水素結合形成と崩壊の間で交換がみられる状態」から考えると,その変異体の His64 は ii)の状態に属する.一方,その変異体の His64 の K_T を野生型のそれ(=1.0)と比較すると,変異体 His64 はより i)の状態に近い.そのため,変異体 His64 は水素結合を形成しない Nδ1 窒素の割合が比較的高いと考えられる.これを His64-$(H_2O)_n$-Zn^{2+}領域において考えると,水素結合を形成しない間に,外部の溶媒とこの水素結合ネットワークの間で基質である水の交換がより容易となることが予想され,このことが酵素活性の上昇の要因となっていることがほのめかされる. 変異体の pKa 値は, 野生型のそれ(=7.2-7.3)と比較すると,0.4 低い結果であった. 現在さらに Y7F 変異体の詳細な理論的解析を進める.これらは量子化学的計算として大きな系に属するため,ハイスペックかつ高機能な計算機マシンが必要とされた.本学 SX-9 マシンによって国際的にハイレベルな競争にあるこの領域に対し本研究を遂行することが可能である.

講演リスト(1-3) ・ 発表論文(4-6)：

1. 量子化学的手法を用いた炭酸脱水酵素の触媒反応におけるプロトントランスファのメカニズムの解析, 島原秀登, 杉森公一, 吉田卓也, 長尾秀実, 中沢隆, 大久保忠恭, 小林祐次, 第 6 回日本炭酸脱水酵素研究会学術集会, 東京, 2010/09/12

2. 炭酸脱水酵素の触媒反応におけるプロトントランスファの量子化学的解析のためのモデル－互変異性が変化する状態にある His64 のスウィング運動の関与 について, 島原秀登, 杉森公一, 吉田卓也, 長尾秀実, 中沢隆, 大久保忠恭, 小林祐次, 第 33 回日本分子生物学会年会・第 83 回日本生化学会大会合同年会, 神戸, 2010.12.07-10

3. Theoretical investigation of primary and secondary H/D isotope chemical shift, K. Sugimori, H. Kawabe, H. Shimahara, The 51st Sanibel Symposium, Georgia, USA, 2011.02.25-03.01

4. 北陸の研究たより, 島原秀登, HIAC NEWS, 79, 14, 2010/08

5. 支部だより-研究テーマ-, 島原秀登, 生物物理学会誌, 51, 1, 48-9, 2011/01

6. Assembly states of the nucleosome assembly protein 1 (NAP-1) revealed by sedimentation velocity and non-denaturing mass spectrometry. Noda M, Uchiyama S, McKay AR, Morimoto A, Misawa S, Yoshida A, Shimahara H, Takinowaki H, Nakamura S, Kobayashi Y, Matsunaga S, Ohkubo T, Robinson CV, Fukui K. 2011 Feb 24. [Epub ahead of print] 査読あり

高精度擬ポテンシャル法の開発と表面系への応用

川井 弘之（尾崎研究室）

【緒言】実験データや経験パラメータを用いずに、シュレーディンガー方程式（ディラック方程式）から物性・化学反応予測を行う第一原理計算は、コンピュータの進歩に伴い、ますます重要になってきている。しかし、全電子計算法の巨大な系への適用は未だに困難である。このため、内殻電子の寄与をある有効ポテンシャルで近似するという、擬ポテンシャル法がよく使用されている。我々は、重い原子系に適用可能な高精度ノルム保存型 Vanderbilt 擬ポテンシャルを開発した。従来のノルム保存型擬ポテンシャルは、各 l チャンネルに対して、単一の参照エネルギーを用いるため、重い原子への適用は問題があったが、参照エネルギーを複数取ることができる Vanderbilt のノルム保存型擬ポテンシャル（Morrison-Bylander-Kleinman の擬ポテンシャル）を開発することで、この問題を解決した。また、新しく開発した高精度ノルム保存型擬ポテンシャルを用いて、最近 Fleurence らによって見いだされた、ZrB_2 表面上のグラファイト様 Si 単一層（silicene）の電子構造計算を行い、XPS による実験スペクトルとの比較検討を行った。

【計算】はじめに、Si 原子の 2p 軌道を考慮しない MBK 擬ポテンシャルを用いて Si on ZrB_2 の構造最適化計算を行った結果、Fig. 1 のような構造が得られた。次に、Si 原子の 2p 軌道を考慮した MBK 擬ポテンシャルを作成し、この擬ポテンシャルによって、Fig. 1 の構造を用いた Si on ZrB_2 の一点計算を行った。そして、この計算結果と XPS の実験結果を比較した。

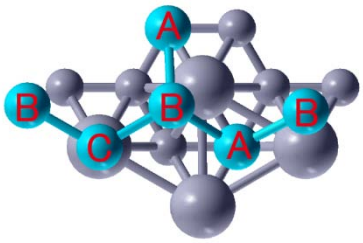
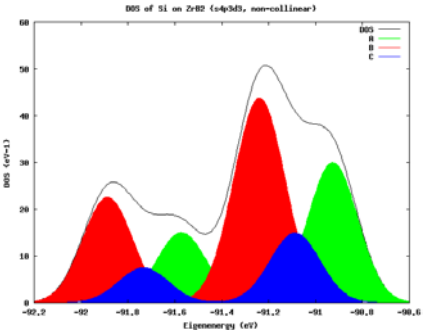
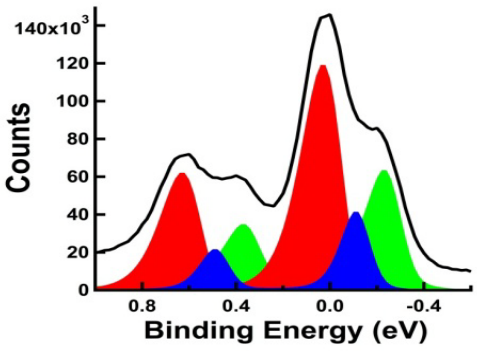


Fig. 1　Si on ZrB_2 の構造

【結果】計算結果の状態密度（Density Of States, DOS）のプロットを描いたところ、hollow (A), bridge (B), on-top (C)の各位置に相当するピークが得られ（Fig. 2）、Fleurence らによる XPS の実験結果（Fig. 3）と一致していることが示された。





【Keywords】第一原理計算、擬ポテンシャル、グラファイト様 Si、silicene

【計算機利用法】Cray XT5 を MPI と OpenMP のハイブリッド並列（計 128 並列）で利用

研究業績

1. 高精度擬ポテンシャル法の開発と表面系への応用，修士論文
2. Electronic structure calculations of a Si mono layer on ZrB_2 by an accurate norm-conserving pseudopotential method, H. Kawai, T. Ozaki, A. Fleurence, and Y. Yamada-Takamura, The 13th Asian Workshop on First-Principles Electronic Structure Calculations, POSTECH, Pohang, Korea, November 2 , 2010　ポスター発表
3. ZrB_2 上の Si 単一層の第一原理電子状態計算，川井弘之，尾崎泰助，Antoine Fleurence，Rainer Friedlein，高村（山田）由起子，日本物理学会第 66 回年次大会，新潟大学五十嵐キャンパス，2011.03.25

第一原理計算プログラム OpenMX の次期バージョン基底関数の検証

日産自動車株式会社 総合研究所　大脇 創

使用計算機：XT5
計算規模：32～128 プロセッサ，計算時間：最大で 30 日程度

【概　要】第一原理計算プログラム OpenMX[1]で用いられる局在基底関数の現行版および次期改良版[2]について，分子間相互作用エネルギー値の計算精度の観点から詳細な検証を行なった。改良版では平面波基底法の計算精度を double-zeta + single-polarization で確保できることを確認した。

【背　景】本研究では Li イオン二次電池（LIB）の界面における電気化学反応を解明することを目的に，OpenMX に基づく第一原理分子動力学シミュレーションを行っている。反応場としての界面構造を詳細に検討するためには，表面と電解質分子，そして電解質分子間の相互作用が高精度に扱われなければならない。OpenMX で用いられている基底関数は（擬原子模型で構築された）局在基底関数であるため，BSSE（Basis Set Superposition Error）の克服がシミュレーション精度を高める重要なポイントとなる。本研究では，OpenMX の基底関数の精度について，分子間相互作用エネルギーの観点から検討した。

【計算モデル・計算条件】LIB 用有機電解質の一つであるプロピレンカーボネート（PC）分子のダイマー構造（Figure 1）における分子間相互作用エネルギー値を検証対象とした。PC 分子のダイマー構造は CASTEP ver. 3.4（平面波基底法＋ウルトラソフト擬ポテンシャル法）で最適化し，一連の検証計算はその構造の下で行った。交換相関汎関数は GGA-PBE とした。

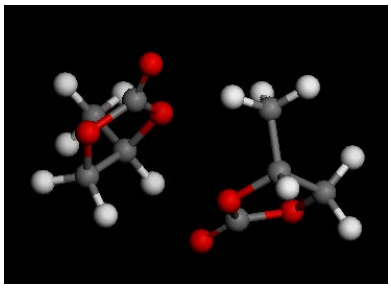


Figure 1. PC 分子ダイマーの構造.

【計算結果および考察】

１．平面波基底法での相互作用エネルギー値：CASTEP での計算によると，PC 分子間の引力的相互作用エネルギー値は −4.0 kcal/mol であった。この平面波基底法で得られた値を基準に，以後の検証を行うことにした。

２．従来版の基底関数についての精度検証：各価電子軌道および分極軌道に対する Primitive Orbital（PO）線形展開中の PO 数を変えた場合の，PC ダイマーにおける相互作用エネルギー値の変化を計算し，Table 1 にまとめた。ここでは，分極軌道として炭素，酸素原子に対して d 軌道（一部 d 軌道と f 軌道），水素原子に対して p 軌道（一部 p 軌道と d 軌道）を付加して検証計算を行った。尚，表中の表記について，例えば“s2”とは，価電子軌道の 2s 軌道を２つの PO の線形結合で展開して基底関数を構築したことを意味する。

Table 1 によると，平面波基底で求められた相互作用エネルギー値 −4.0 kcal/mol に対し，いわゆる chemical accuracy の 1 kcal/mol の範囲内に入るのは，各軌道における PO 数が 7 以上の場合（C,O: s7p7d7 / H: s7p7）である。しかしながらこの場合，PC 分子 1 個当たりの PO 数は 509 にもなり，これでは数十～数百分

Table 1. 各価電子軌道および分極軌道に対する PO 線形展開中の PO 数を変えた場合の，PC ダイマーにおける相互作用エネルギー値の変化（OpenMX ver. 3.5，GGA-PBE，基底関数のカットオフ半径 = 7.0 Bohr）.

Basis set (C, O / H)	Interaction energy (kcal/mol)
s1p1d1 / s1p1	–47.6397
s2p2d2 / s2p2	–14.5614
s3p3d3 / s3p3	–9.3178
s4p4d4 / s4p4	–7.0822
s5p5d5 / s5p5	–5.5954
s6p6d6 / s6p6	–5.0649
s7p7d7 / s7p7	–4.8852
s8p8d8 / s8p8	–4.7987
s9p9d9 / s9p9	–4.7162
s10p10d10 / s10p10	–4.6633
s15p15d15 / s15p15	–4.6197
s20p20d20 / s20p20	–4.6112
s20p20d20f10 / s20p20d10	–4.4786

子を含む大規模モデル系で計算を行うには非常に高い計算負荷が発生し，計算の実行が困難となる。また，分極関数として C，O 原子に d 軌道，水素原子に p 軌道を付加させる範囲では，相互作用エネルギー値はおよそ −4.6 kcal/mol が収束値となっている。しかし，高い方位量子数の分極関数（f および d 軌道）を加えると更に −4.4 kcal/mol 台に低下することから，今回の検証範囲では十分な相互作用エネルギー値の収束は得られないことが分かった。

３．次期改良版の基底関数についての精度検証：上述の検証結果を受け，OpenMX の開発者である尾崎泰助准教授は OpenMX の新しいデータベース（擬ポテンシャルおよび基底関数）を構築した[2]。その改良版の基底関数に基づいた相互作用エネルギー値の計算について，まずカットオフエネルギー値に対する依存性について検証したところ（s2p2d2 / s2p2），150 Ry から −4.5 kcal/mol 辺りで収束していることが確認できた（尚，190 Ry の場合で若干特異的にエネルギー値が小さくなっているが, 収束値から 1 kcal/mol 未満の差だったため，このカットオフエネルギー値についても一連の収束傾向の範囲内にあるとみなした）。

次に，PC 分子ダイマー内における相互作用エネルギー値の PO 構成依存性を Table 2 に示した。従来版の基底関数とは大きく異なり，少ない PO 数で精度の高い相互作用計算が可能となっていることが分かる。s2p2d1/s2p1 で平面波基底法でのエネルギー値 −4.0 kcal/mol から chemical accuracy の範囲内に達し，s2p2d2/s2p2 で約 −4.0 kcal/mol となっている。更に PO 数を増加させた場合でも，−4.0 kcal/mol 辺りで収束している。

Table 2. 改良版の基底関数を用いて計算された，PC 分子ダイマー内における相互作用エネルギー値の PO 構成依存性（基底関数のカットオフ半径 = 7.0 Bohr）.

Basis set (C, O / H)	Interaction energy (kcal/mol)
s2p2 / s2	–6.7297
s2p2d1 / s2p1	–4.7791
s2p2d2 / s2p2	–3.9586
s3p3d2 / s3p2	–3.9899
s3p3d3 / s3p3	–4.0202
s4p4d4 / s4p4	–3.9409
s8p8d8 / s8p8	–3.8058
s10p10d10 / s10p10	–3.8276
s15p15d15 / s15p15	–3.8603

更に OpenMX の機能の一つである基底関数最適化に基づく基底関数の高精度化を件とした。基底関数最適化計算は PC 単分子で行い，カットオフエネルギー値は 170 Ry とし，s2p2d1/s2p1 構造の最適化基底関数（double zeta + single polarization）を得ることにした。

Table 3 に各 contraction 長で最適化された基底関数での相互作用エネルギー値をまとめた（表中の表記で，例えば“s6>2”とは，s 軌道の最適化の際に６つの PO から２つの PO に縮約するという意味）。この表によると，今回検討した contraction 長の範囲で相互作用エネルギー値の良好な改善が確認された。いずれの場合でも，平面波基底法での計算値 −4.0 kcal/mol との差が約 0.1 kcal/mol となっており，平面波基底法の精度が十分再現されていることが分かる。

【まとめ】今回の OpenMX に関する検討により，比較的規模の小さい double zeta + single polarization の基底関数で平面波基底レベルの計算精度で分子間相互作用エネルギー値の計算が可能であることが確認された。これにより，今後は LIB 系における界面の第一原理分子動力学シミュレーションを，高精度化と大規模化を両立させながら実行することが出来るようになった。

Table 3. 改良された基底関数を基に各 contraction 長で最適化された基底関数での PC 分子ダイマー中の相互作用エネルギー値（基底関数のカットオフ半径 = 7.0 Bohr, 基底関数最適化計算の収束条件 : Gradient Norm ≤ 10^{-6} (Hartree/borh)^2）.

Contraction (C, O / H)	Interaction energy (kcal/mol)
s6>2p6>2d6>1 / s6>2p6>1	–4.0366
s8>2p8>2d8>1 / s8>2p8>1	–4.1475
s10>2p10>2d10>1 / s10>2p10>1	–4.1836

[1] http://www.openmx-square.org/
[2] 尾崎, 川井, 大脇, H. Kim, J. Yu, 日本物理学会 第 66 回年次大会（28aPS-20）.

不均一系 Ziegler-Natta オレフィン重合における活性点のフレキシビリティ

（マテリアルサイエンス研究科 助教） 谷池 俊明

概要

$MgCl_2$ 担持型 Ziegler-Natta 触媒はプロピレンの立体・位置選択的重合に非常に高い活性を示す工業触媒であり、触媒特性発現の核である活性 Ti 種の状態やドナーと総称されるルイス塩基化合物による触媒性能向上機構に関する研究が精力的に行なわれてきた。これらの議論の中心は、担体である $MgCl_2$ の{110}及び{100}平坦表面上での Ti 種やドナーの吸着状態に基づくものであり、中でも Corradini 等は、「ドナーは $MgCl_2$ {110}表面に選択的に吸着し、その上での非立体特異的な Ti 単核種の生成を阻害し、{100}表面上の高イソ立体特異的 Ti 二核種の存在比率を増加させることで触媒の立体特異性を大幅に向上させる」という著名な提案を行なった[1]。これに反して、近年我々は密度汎関数法を用いた網羅的な検討によって、①ドナーと Ti 種は{110}、{100}面に非選択的に吸着すること、②{110}面上でのドナーと Ti 種の共吸着が{100}面上でのそれと比較して、エネルギー・立体・電子的に有利であること、③従来から立体特異的とされてきた{100}面上の Ti 二核種が不安定かつ非立体特異的であること、④{110}面上におけるドナーと Ti 種の共吸着構造は、ドナーによる立体・位置特異性の向上等の多くの実験結果を統一的に説明し得ること等を明らかにした（Figure 1）[2]。

本研究では $MgCl_2$ {110}・{100}平坦表面上での上述の議論を、原子レベルステップを考慮した非平坦表面に展開し、最適活性点構造と実触媒系の表面構造を明らかにすることを試みた。

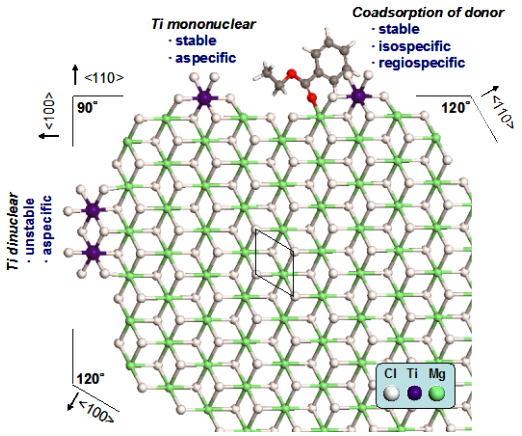


Figure 1. Summary of active sites performance on flat $MgCl_2$ surfaces

計算方法

密度汎関数計算には DMol3 を用い、主に Altix4700 で実行した。交換相関汎関数は PBE とし、基底関数には DNP 及び有効殻ポテンシャルを用いた。$MgCl_2$ {110}・{100}表面はスラブ法による無限遠表面として表し、{110}表面には 6 原子層の p(2x2)のユニットセルを、{100}表面には 14 原子層の p(4x1)のユニットセルを使用した。このような{110}・{100}平坦表面を出発点として、$MgCl_2$ 単位の加減によって各表面上に原子レベルステップ構造をモデリングした（Figure 2）。各活性点のプロピレン重合性能の評価には、気相プロピレンの E を零点とした見かけの活性障壁（ΔE_{ap}）を用い、最安定な *re*・*si* 面挿入間の E 差を立体特異性（$\Delta\Delta E_{stereo}$）、最安定な 1,2・2,1 挿入間の E 差を位置特異性（$\Delta\Delta E_{regio}$）とした。

成果

Figure 2**a-c** に $MgCl_2$ {110}・{100}平坦表面上における代表的な活性点構造を示した。Table 1 に示されるように、ドナーの存在しない**a**,**c** は非立体特異的である。ドナー（安息香酸エチル）が共吸着した活性種 **b** は、高イソ立体特異性を発現し、且つ、**a**,**c** と比較して高い位置特異性を有する。このように、**b** のような{110}平坦表面上における共吸着構造はドナーの添加効果を正確に描写可能である[2]。次に、平坦表面に $MgCl_2$ 単位を加えて原子レベルステップ・テラスを作成した。このような原子レベルステップ・テラスに Ti 種やドナーは非常に強く吸着する。例として、{100}面上の{110}ステップに Ti 単核種が吸着した活性種 **d**、さらにドナー（1,3-ジエーテル）が共吸着した活性種 **e** を Figure 2 に示した。活性種 **d** は、他のドナーを有さない種と同様に特異性に乏しいが、原子レベルステップ上の共吸着構造 **e** は平坦表面における共吸着構造 **b** と同様に高い立体・位置特異性を示し、**b** よりもずっと低い ΔE_{ap} を持つ。すなわち、原子レベルステップにおける共吸着構造は非常に高い潜在活性を有する。このような原子レベルステップ・テラスは $MgCl_2$ 担体には存在しないが、実触媒表面の大半を占めていると考えられる（Figure 2**f**）[3]。

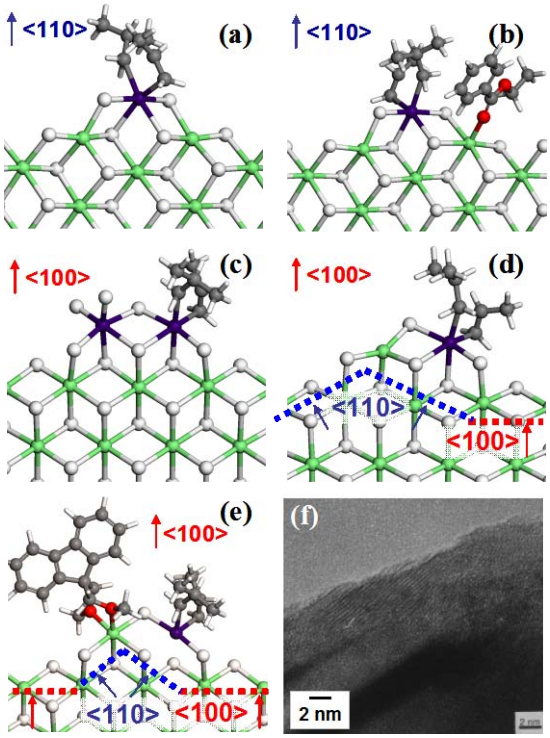


Figure 2. (a) Ti mononuclear on {110}, (b) coadsorption with ethylbenzoate on {110}, (c) Ti dinuclear on {100}, (d) Ti mononuclear at a {110} atomic step, (e) coadsorption with 1,3-diether at a {110} atomic step, and (f) a surface structure of coground $TiCl_4/MgCl_2$ catalyst [3]

Table 1. Active sites performance

Structure	ΔE_{ap} [1)]	$\Delta\Delta E_{stereo}$ [1)]	$\Delta\Delta E_{regio}$ [1)]
a	7.4	0.2	1.3
b	7.9	2.3	2.1
c	1.7	0.1	1.7
d	2.6	0.4	1.3
e	3.4	3.4	2.7

[1)] In kcal/mol.

結論

本研究では、$TiCl_4$ やドナーが触媒調製過程で生じた原子レベルステップ・テラスに非常に強く吸着することで構造を安定化し、自己制御的に安定かつ高性能な活性種を与えていることを初めて発見した。

文献

[1] Busico, V.; Corradini, P.; De Martino, L.; Proto, A.; Savino, V. *Makromol. Chem.* **1985**, *186*, 1279.

[2] Taniike, T.; Terano, M. *Macromol. Rapid. Commun.* **2007**, *28*, 1918; *Next-Generation Polyolefins* **2008**, *2*, 60.

[3] Mori, H.; Sawada, M.; Higuchi, T.; Hasebe, K.; Otsuka, N.; Terano, M. *Macromol. Rapid. Commun.* **1999**, *20*, 245.

発表論文リスト

1. “不均一系 Ziegler-Natta 触媒研究最前線”, 谷池 俊明, 触媒, 2010, Vol. 52, No. 4, 291.
2. “次世代の汎用性ポリオレフィンを支える触媒技術”, 谷池 俊明, 寺野 稔, 高分子, 2010, Vol. 59, No. 11, 849-852.
3. “不均一系 Ziegler-Natta 触媒の性能発現機構に関する密度汎関数計算”, 谷池 俊明, 寺野 稔, 2010, Vol. 52, No. 8, 547-552.

主な国際発表

1. “A Combined Experimental and Theoretical Study for Propylene Polymerization Mechanism with Heterogeneous Ziegler-Natta Catalysts”, Toshiaki Taniike, Minoru Terano, FHI-CRC Joint International Symposium, Sapporo, Japan, Apr. 1-2, 2010, *Invited Lecture.*
2. “Coadsorption Model for the First-Principle Description of the Role of Donors in Ziegler-Natta Propylene Polymerization”, T. Taniike, M. Terano, 3rd Blue Sky Conference on Catalytic Olefin Polymerization, Sorrento, Italy, Jun. 20-23, 2010, *Poster*.
3. “Molecular Genesis of Stereospecificity in Heterogeneous Ziegler-Natta Propylene Polymerization”, T. Taniike, M. Terano, 7th LSP/JAIST International Colloquium on Heterogeneous Ziegler-Natta Catalysts, Sorrento, Italy, Jun. 24-25, 2010, Keynote Lecture.
4. “First Principle Investigation of the active site structures on the Phillips-type Catalyst”, K. Tonosaki, T. Taniike, M. Terano, 7th LSP/JAIST International Colloquium on Heterogeneous Ziegler-Natta Catalysts, Sorrento, Italy, Jun. 24-25, 2010, *Oral*.
5. “First-principle description of Ziegler-Natta propylene polymerization: active site and polymerization mechanism”, Toshiaki Taniike and Minoru Terano, Computer Simulation Studies of Soft Materials, Ishikawa, Japan, July. 30, 2010, *Invited Lecture.*
6. “Effects of Exchange-Correlation Functionals on Ziegler-Natta Propylene Polymerization”, T. Taniike, M. Terano, International Workshop on Computational Approach for Heterogeneous Ziegler-Natta Catalysts, Ishikawa, Japan, Feb. 24-25, 2011, *Invited Lecture.*

全学共用計算サーバ・並列計算機利用レポート

尾崎泰助　先端融合領域研究院
使用計算サーバ：xt5
計算規模: 1〜128 プロセッサー, 計算時間は最大で 7 日程度

研究の概要；『オーダーN 大規模密度汎関数法コード:OpenMX の開発と応用』

密度汎関数理論は金属、半導体、生体分子等の広範囲の物質群に適用可能な汎用性の高い手法であるが、その計算時間は系のサイズ(基底関数もしくは原子数)の三乗に比例しており、ナノメーターサイズの系を取り扱うことは実際上、困難である。我々は大規模系への密度汎関数理論の適用限界を拡張するために計算量が原子数に単に比例したオーダーN 法と呼ばれる新しい計算手法を開発している。開発された新しい計算手法はプログラムパッケージ:OpenMX (http://www.openmx-square.org/)[1] に組み込み、GNU-GPL の規約の下で無償公開している。本年度の研究内容を以下に要約する。

(1) **大規模密度汎関数法のための近似のない高並列化低次スケーリング法の開発**

Car-Parrinello 法の発展を契機に、第一原理密度汎関数計算の大規模系への適用限界を拡張するために非常に多くの対角化計算手法がこれまでに開発されている。これらの対角化手法は(i) 共役勾配法や Davidson 法などの繰り返し計算に基づくものと、(ii) 近似を導入して計算量を O(N)に低減したオーダーN 法に大別される。前者においては与えられた Kohn-Sham ハミルトニアンに対して近似を導入することなしに、固有状態の計算を行うが、占有軌道間の直交条件を保持する計算に対して $O(N^3)$の計算量が必要となる。その中間に位置する $O(N^{2\sim})$法の開発はこれまであまり行われていないが、近似を導入することなしに計算オーダーが低減できるならば、$O(N^3)$法に代わる手法として興味の持たれる所である。今回、擬原子局在基底や有限要素法などの局在基底関数の場合に対して、近似を導入することなしに低い計算オーダーを持った新しい対角化手法を開発した[2]。アイデアは留数積分[3]によって直接、密度行列を計算するというものであり、収束に必要な留数の数が系のサイズに依存しないことを示すことができる。主要な計算は疎行列の逆行列計算(グリーン関数の計算)となり、ここで計算オーダーが支配される。我々は逆行列要素を計算するために、行列の nested dissection(ND)法と Shur 補元を用いて漸化式を導出し、これを用いて必要な逆行列要素を求める新しい計算手法を開発した。この漸化式に基づく逆行列計算法の計算オーダーは 3 次元系で $O(N^{7/3})$、2 次元系で $O(N^2)$、1 次元系で $O(N(\log_2 N)^2)$であり、近似を導入することなしに計算オーダーを低減することが可能となった。

(2) **原子 Hartree-Fock 及び密度汎関数法計算のための高精度有限要素法の開発**

交換・相関汎関数を開発する基盤技術として、有限要素法に基づく原子の高精度電子状態計算手法を開発した[4]。密度汎関数理論における局所密度近似(LDA)を改良する一つの手段は、波動関数理論を利用して、近似的な密度汎関数では計算精度が確保出来ないエネルギー成分を順次、波動関数もしくは密度行列の汎関数として取り込むことである。通常の Kohn-Sham(KS)法では運動エネルギーの大部分を KS 軌道から算出し、残りのエネルギー成分を密度汎関数として取り扱う。この改良の次のステップは(1) 交換エネルギー、(2) 一部分の相関エネルギーを波動関数もしくは密度行列の汎関数として順次、取り込んでいくことである。

この様な近似交換・相関汎関数を開発する上で重要な事は数値的な誤差と手法自体から生じる誤差を明確に区別することであり、基底関数や数値積分から生じる誤差が実質的にゼロである手法を用いることが望ましい。我々は近似交換・相関汎関数の検証の舞台としてまず原子系を選択し、Hermite スプライン関数を用いた高精度電子状態計算手法を開発した。LDA 及び Hartree-Fock 近似に本手法を適用し、全エネルギーの誤差を virial 定理から見積った所、Z=1-103 の全ての元素に対して 10^{-7}Hartree 以下の絶対誤差で全エネルギーが計算できることが分かった。

発表論文等

[1] http://www.openmx-square.org/

[2] “Continued fraction representation of the Fermi-Dirac function for large-scale electronic structure calculations”, T. Ozaki, Phys. Rev. B 75, 035123 (2007).

[3] “Efficient low-order scaling method for large-scale electronic structure calculations with localized basis functions”, T. Ozaki, Phys. Rev. B 82, 075131 (2011).

[4] “Accurate finite element method for atomic calculations based on density functional theory and Hartree-Fock method”, T. Ozaki and M. Toyoda, Comp. Phys. Comm. 182, 1245 (2011).

一列に穴が開いたグラフェンの電気伝導シミュレーション

株式会社富士通研究所　基盤技術研究所
實宝秀幸

使用計算サーバ ： Cray-XT5
計算規模 ： 1～128 プロセッサー（計算時間は最大 2 日程度）

【背景】

グラフェンは高移動度を持つことで知られるが半金属であるため、電子デバイス応用するためにはバンドギャップの制御が重要である。昨年、ナノスケールの穴をメッシュ状に開けた構造（図 1(左)）がバンドギャップ持つことが実験的に示された[1,2]。しかし、チャネル長が数十ナノメートル程度のデバイスでは穴が一列になると考えられ（図 1(右)）、その場合の電気伝導特性についての知見が求められている。

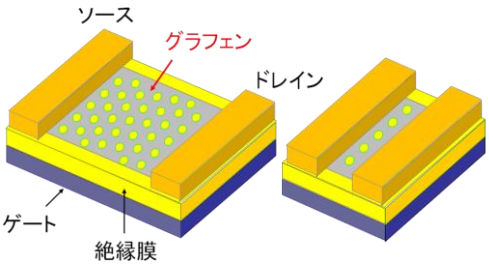


図 1. メッシュ状(左)および一列(右)に穴が開いたグラフェンをチャネルとしたトランジスタの模式図。

【結果と議論】

図 2 に示した原子構造モデルを考え、一列に穴が開いたグラフェンについて、非平衡グリーン関数法に基づいた第一原理シミュレーションを行った。用いた計算コードは OpenMX である。得られた電流電圧特性を図 3 に示す。このモデルでは両端に半無限のグラフェン電極が接続されているため、伝導方向の閉じ込め効果が無くなって、電極間にバイアスをかけても電流の流れない伝導ギャップが非常に小さくなることが推測される。実際に構造(a)のギャップは 0.1eV 以下と非常に小さい。しかし、構造(b)では 0.5eV 程度のギャップが存在する。構造(b)の穴の端はジグザグ構造となっており、その近傍に電子が局在したエッジ状態と呼ばれる状態を広いエネルギー範囲に作る。これらが伝導に寄与しないため、伝導ギャップが大きくなっていることが分かった。

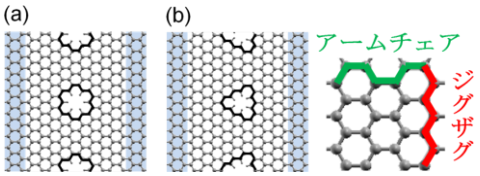


図 2. (a)ジグザグ端がない穴の列と、(b)ジグザグ端がある穴の列を持つグラフェン構造。穴の端は黒太線で強調されている。青い領域はグラフェン電極である。右図はグラフェンの 2 つの原子配置の呼称の説明。

このように、一列の穴でも伝導ギャップを持つことが分かり、ナノスケールデバイスへの応用の可能性を示すことができた。ただし、伝導ギャップを決定する上でジグザグ端が重要な役割を果たすことも示され、デバイス応用には穴の端の構造制御が重要であると考えられる。

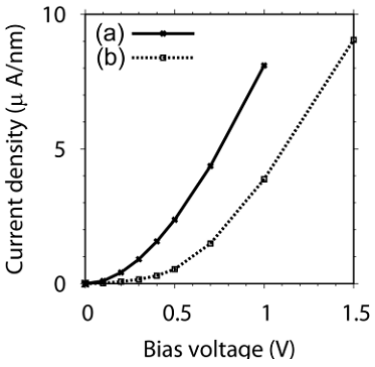


図 3. 構造(a)(b)の電流電圧特性。

【参考文献】

[1] J. Bai *et al.*, Nature Nanotech. 5, 190 (2010).
[2] X. Liang et al., Nano Lett. 10, 2454 (2010).

【学会発表】

1) H. Jippo, M. Ohfuchi and C. Kaneta, “First-principles study on periodic nanoholes in a graphene sheet”, Recent Advances in Graphene and Related Materials, Singapore, August 1-6, 2010

2) H. Jippo, M. Ohfuchi and C. Kaneta, “Theoretical study on graphene sheets with periodic nanoholes”, 第 39 回フラーレン・ナノチューブ総合シンポジウム, 京都大学, 2010 年 9 月 5-7 日
3) 實宝、大淵、金田、“周期ナノホールを持つグラフェンシートの第一原理計算”、長崎大学、2010 年秋季 第 71 回 応用物理学会学術講演会
4) H. Jippo, M. Ohfuchi and C. Kaneta, “Theoretical study on graphene sheets with periodic nanoholes”, 23rd International Microprocesses and Nanotechnology Conference (MNC), Fukuoka, Japan, November 9-12, 2010
5) H. Jippo, M. Ohfuchi and C. Kaneta, “Significance of zigzag edges in electon transport properties of graphene sheets with periodic nanoholes”, 第 40 回フラーレン・ナノチューブ総合シンポジウム, 名城大学, 2011 年 3 月 8-10 日

Adsorption of Carbon-based nanostructures on metallic substrates: a DFT approach to large-scale calculations.

Giulio Biddau, Ruben Perez, Taisuke Ozaki
Departamento de Fisica Teorica de la Materia Condensada, Universidad Autonoma de Madrid, Spain
Research Center for Integrated Science, Japan Advanced Institute of Science and Technology, Nomi City, Japan

XT5
average daily use of 32 CPU

In this work we have first studied the adsorption of large hydrocarbons on highly reactive metallic substrates, in particular, the Pt(111) surface. We have characterized how the adsorption of technologically important large molecules, like Polyciclic aromatic hydrocarbons, can be understood in terms of the behavior of one of their subunits (a benzene ring) in the same substrate, both structurally and electronically. Based on this analysis, we have developed an approximate and efficient method to explore the phase space to save computational time in large scale calculations.
In a second stage, we analyzed the extremely challenging de-hydrogenation process of these molecules on Pt(111). Through molecular dynamics simulations we obtained an outstanding agreement with experimental results and provide an otherwise impossible insight in this important process to achieve carbon nanostructure formation.
In the third part of the work, we have considered the functionalization of the hydrocarbons. In this way we have seen how site specific doping influences many different characteristics, among them the structural, the electrical properties and, nonetheless, the de- hydrogenation and cyclization process. Our finding show that site specific doping is extremely relevant for this new branch of synthesis of molecules.
In the last part of the work we have considered the adsorption of fullerenes on Pt(111), analyzing the mutual fullerene interaction in various reconstructions and considering an exhaustive description of various adsorption configurations in the many possible orientations and their energetic convergence respect to the number of substrate layers.
This topic has been approached with both LDA and GGA approximations and revealed that large scale calculations are required to obtain an appropriate energetic description, nonetheless the study shows that structurally approximations may be done to save computational time.
Functionalization of the fullerene through nitrogen doping has been considered and the magnetic properties studied for the newly obtained triazafullerene.
To conclude we examine effect of substrate defects such as vacancy in the adsorption of the (hetero)fullerenes and obtained a very good agreement with experimental findings.

First-principles study of oxygen reduction reaction and hydrogen oxidation reaction with carbon alloy catalysts

Sheng-Feng Huang, Research Center for Integrated Science

Nitrogen doped carbon-alloy catalysts (CAC) are already proved to be sufficient alternative material used in polymer electrolyte fuel cell (PEFC), to substitute platinum as catalysts in the oxygen reduction reaction (ORR) at cathode. A series of first principles research works have been done for understanding the reactivity and stability of CAC by OpenMX code with pseudoatomic orbital basis functions, and the reaction mechanism of ORR by CPMD code with first principles molecular dynamics simulations. The results are consistent and could explain the phenomena observed in the experiments.

Furthermore, the carbon K-edge X-ray absorption spectra (XAS) are also simulated by CP2k code with Gaussian and augmented plane wave basis function, and half-core hole Slater transition potential method. The various edge terminations by hydrogen are considered in the calculations of graphene nanoribbon systems, which give important information to understand and predict transport and catalysis functions of CAC. It shows that different edge terminations significantly influence the binding energy of 1s core-level of carbon atoms in the vicinity of edges, because of the change of chemical bonding and localized states. By analyzing the XAS simulations with systems in various edge terminations, it gives hints to realize the structure of actual samples.

In addition to the ORR at the cathode of PEFC, Pt-free catalysts in the hydrogen oxidation reaction (HOR) at the anode would be the next aim. Platinum based catalysts show extremely high efficiency in HOR, it might be difficult to find compatible materials. According to the previous results, CAC with boron doping might be a sufficient candidate. Studying of HOR mechanism and calculations of B-doped CAC are now in progress.

Published paper:
Phys. Rev. B **80**, 235410 (2009)
J. Phys. Chem. C **115**, 5392 (2011)

Used machine:
Cray XT5, pcc

Ab-initio study of cyclopentasilane and liquid cyclopentasilane

Pham Tien Lam
School of Materials Science
Used machine: Cray-XT5 and Altixs.

In recent years, liquid processes for fabricating electronic devices have received considerable attention. Compared to the conventional processes, the liquid processes improve the material utilization efficiency, simplify the processing procedures, and reduce the size and cost of the manufacturing apparatus. The use of the liquid processes enables us produce large-scale electronic circuits and open up new applications. The preparation of functional solutions which are stable solutions containing materials for a target device is the most important step in the liquid processes. For the production of electronic devices among a variety of silicon compounds, cyclopentasilane (Si_5H_{10} - CPS) is the most suitable candidate for a source material of functional solutions because it has the ability to undergo ring-opening polymerization and to transform to high purity Si. In addition, CPS can also act as a solvent of polysilanes.

We studies of the physical properties of the liquid cyclopentasilane. We focused on interactions between CPS molecules and the bonding nature in the CPS solution. The investigations were mainly based on ab-initio. Ab-initio molecular dynamics simulations were performed by Car- Parrinello method. The study has showed that the Si-H bonds in a CPS molecule that point to the center of the ring of the others CPS molecule play a significant role in the interaction between CPS. This interaction results in the formation of special bonds between CPS molecules which resemble “hydrogen bonds”. Similarly to those in the water system these hydrogen bonds cause the red shift of IR absorption peaks corresponding to the Si-H stretch vibration. The formation of hydrogen bonds in liquid CPS was further confirmed by ab-initio molecular dynamics simulations. The radial distribution functions of the center of mass of CPS molecules and H atoms, $g_{MC\text{-}H}(r)$, and Si atoms $g_{MC\text{-}Si}(r)$ were analyzed. $g_{MC\text{-}H}(r)$ shows a peak at about 2.0 Å that is consistent wiht the hydrogen atom which involves in a hydrogen bond, while $g_{MC\text{-}Si}(r)$ appears with a peak at about 3.3 Å of the Si atom in a hydrogen bond. The results of the present study have provided important insights into the interaction of CPS molecules in the liquid phase and the complex polymerization processes in the CPS liquid system.

We used density functional calculations on Altixs machine with 8 cores / job by using Dmol3 code. Ab-initio molecular dynamics simulations were performed on Cray-XT5 with 128 cores / job by using CPMD code.

Publication:
Ab-initio study of the intermolecular interaction of cyclopentasilane and liquid cyclopentasilane (To be submitted)

Influence of surface ligands on the electronic structure of Fe-Pt clusters

Trinh Thang Thuy
Maenosono Lab., School of Materials Science
Machine in use: Cray XT5
Typical computational size: 8 node/32 core

Abstract. The geometrical and electronic structures of a chemically disordered face-centered-cubic- (*fcc*) FePt cluster capped with various organic ligands, including propanoic acid, propylamine, and propanethiol, were investigated by means of density functional theory (DFT) calculations within a generalized gradient approximation (GGA). Detailed analysis of the electronic structure revealed that (1) Fe atoms are the favored adsorption sites of the ligands on the surface of the FePt cluster; however, for propanethiol, adsorption can also occur at Pt sites. (2) The spin magnetic moment of Fe atoms at adsorption sites in the clusters containing adsorbed ligands decreases slightly compared to that in the bare cluster on the adsorption of the ligand, and it does not depend on the length of hydrocarbon chain of the ligand. The decrease in the magnetic moment originates from the interplay between the strong hybridization of the majority *d* states of Fe atoms with majority *p* states of O, N, and S atoms and the electron transfer between the ligands and Fe atoms on the surface of the clusters involving *d*, *p*, and *s* states of the Fe atoms, as well as from the high symmetry of the surface Fe atoms on adsorption of a ligand.

Publication. Thuy T Trinh, Taisuke Ozaki, and Shinya Maenosono, Influence of surface ligands on the electronic structure of Fe-Pt clusters: A density functional theory study, *Phys. Rev. B,* 83, 104413 (2011).

Investigation of two hydrogen terminated graphene doped with Nitrogen

Department of Organic and Polymeric Materials, Tokyo Institute of Technology
Xianlong WANG
Using machine XT5

Graphene clusters doped with nitrogen show high oxygen reduction reaction (ORR) activities. [1] Our previous results show that the N doped at the site next to zigzag edge (edge-1 N) can activate the neighboring two edge carbons. However, the N (edge N) located at edge site is the most stable configuration, and edge N may suppress the ORR activity. [2, 3]

As the ratio of two hydrogen terminated edge carbons in graphene clusters will increase with the increasing of the hydrogen partial pressure. Hence, the N doping configurations in different ratio of two hydrogen terminated carbons are analyzed, and the stability of the two hydrogen termination is also studied. Along zigzag edge, the two hydrogen terminated carbons intend to separate each others as far as possible. With the concentration of two hydrogen termination increasing, the energy difference between edge and edge-1 N become smaller and smaller. Because localized state form at the edge-1 site, after the concentration of two hydrogen termination increase to 50%, edge-1 N is the most stable configuration. Interestingly, comparing with CH_4 and C_6H_{12} molecules, the sp^3 bonding energy of C-H bond in the case of two hydrogen terminated edge carbons is small. The second hydrogen can easily be removed by oxygen molecule. After N doping at the edge-1 position, the bonding energy of the second hydrogen will decrease furthermore. This indicate that more powerful catalyst for ORR can be produced in the following way: Firstly, using high hydrogen partial pressure, more two hydrogen terminated edge carbons are created, hence, concentration of edge-1 N is increased, after that, we can expose the samples in air, the second hydrogen will be removed, finally, the activated samples with high edge-1 N concentrations are produced.

References:

[1] J. Ozaki, T. Anahara, N. Kimura and A. Oya, Carbon. **44**, 3358 (2006),

[2] T. Ikeda, M. Boero, S.F. Huang, K. Terakura, M. Oshima and J. Ozaki, J. Phys. Chem. C **112**, 14706 (2008),

[3] S.F. Huang, K. Terakura, T. Ozaki, T. Ikeda, M. Boero, M. Oshima, J. Ozaki and S. Miyata, Phys. Rev. B **80**, 235410 (2009).

The Report on Use of Computing Facilities of JAIST

Department of Organic and Polymeric Materials, Graduate School of Science and Engineering, Tokyo Institute of Technology

Zhufeng HOU

Used machines: Cray xt5 and PC cluster (pcc)

1. Carbon *K*-Edge X-ray Absorption Spectra of Nanographene

X-ray absorption spectroscopy (XAS) is widely used in experiment to characterize the electronic, structural, and bonding information not only about nanographene but also about atoms belonging to functional groups at the surface or at the edge of nanographene. However the assignments for some of the main features in the carbon *K*-edge XAS spectra of graphene-related materials are ambiguous in experiment (the detailed summarization is described in our publication): i) the origin of a shoulder below the π* peak (~285 eV); ii) the origin of the features appearing at ~287 eV and ~288 eV, which are just between the π* peak and σ* peak (~292 eV). In order to complement experiments and straightforwardly provide more convincing assignment for the carbon *K*-edge XAS spectra of nanographene, we have inspected the effects of hydrogen termination of graphene edges in various ratios of mono- to di-hydrogenation.

We have performed the density functional theory (DFT) calculations on the carbon *K*-edge *x*-ray absorption spectroscopy (XAS) spectra of graphene nanostructures including infinite sheet, cluster, and nanoribbons (NRs). To understand the effect of H-termination on the spectra, the chemical shifts of C 1*s* core-level binding energies in NRs with different concentrations of H-termination at the edges are studied. Our results show that H termination at the edges of NRs can significantly affect the binding energy of 1*s* core-level of edge carbon. The 1s core-level of edge C with sp^2 C-H binding (with sp^3 C-H bonding) is shallower (deeper) than that of interior C with sp^2 C-C bonding in NRs. We find that a shoulder appears below the π* peak in the theoretical spectra of NRs with zigzag edges, but its position relative to the π* peak depends on the ratio of 1H-termination to 2H-termination at the zigzag edges. In the theoretical spectra for the cluster and NRs of graphene, two features between the π* peak and σ* peak are found, one of them located around 287 eV originates from the sp^2 C-H bonding in the 1H-terminated edge C, and the other one located around 288 eV is ascribed to the sp^3 C-H bonding in the 2H-terminated edge C. Our results will provide useful information not only to assign the features observed in experiments, but also to experimentally check the electronic structure sensitivity of graphene NRs to the hydrogen partial pressure.

The calculations were performed using the CP2K code and the Quantum-Espresso software package.

Publication:

Zhufeng Hou, Xianlong Wang, Takashi Ikeda, Shen-Feng Huang, Kiyoyuki Terakura, Mauro Boero, Masaharu Oshima, Masa-aki Kakimoto, and Seizo Miyata, *Effect of Hydrogen Termination on Carbon K-Edge X-ray Absorption Spectra of Nanographene*, J. Phys. Chem. C, **115** (13), pp 5392–5403, (2011).

2. First-Principles Simulations of STM Images of Nitrogen-Doped Graphene with Structural Defects

We have performed the first-principles calculations on the nitrogen doping in graphene with different types of simple structural defects (single vacancy, divacancy and its reconstructed configurations, Stone-Wales defect). The simulation of STM images of nitrogen-doped graphene with structural defects was carried out according to the Tersoff-Hamann approximation. The energetically favorable sites of nitrogen dopant in graphene with structural defects and the interaction between two nitrogen dopants are studied. We find that the position of the defect states induced by structural defects can be tuned by the doping site and concentration of nitrogen dopant and hence the simulated STM images of graphene with structural defects are significantly affected by the nitrogen doping.

The calculations were performed using the OPENMX code and the Quantum-Espresso software package.

4. 知識科学分野の計算サーバ利用研究

Modeling the diversity and log-normality in data

Khoat Than and Tu-Bao Ho

Used Machine: Cray XT5, Altix

We investigate two important properties of real data: diversity and log-normality. Log-normality accounts for the fact that data follow the lognormal distribution, whereas diversity measures variations of the attributes in the data. To our knowledge, these two inherent properties have not been paid much attention from the machine learning community, especially from the topic modeling community. In this article, wefill in this gap in the framework of topic modeling. We first investigate whether or not these two properties can be captured by the most well-known Latent Dirichlet Allocation model (LDA), and find that LDA behaves inconsistently with respect to diversity. Particularly, it favors data of low diversity, but works badly on data of high diversity. Then, we argue that these two inherent properties can be captured well by endowing the topic-word distributions in LDA with the lognormal distribution. This treatment leads to a new model, named Dirichlet-lognormal topic model (DLN). Using the lognormal distribution complicates the learning and inference of DLN, compared with those of LDA. Hence, we used variational method, in which model learning and inference are reduced to solving convex optimization problems. Extensive experiments strongly suggest that (1) the predictive power of DLN is consistent with respect to diversity, and that (2) DLN works consistently better than LDA for datasets whose diversity is large, and for datasets which contain many log-normally distributed attributes. Justifications for these results require insights into the used statistical distributions and will be discussed in the article.

This work relates significantly intensive computation. Hence we use some high performance servers of JAIST, for instance, XT5 and Altix. Until this time, because of our programming limitation, we only use interactive interface to work on those machines. Since our programs are not for parallel, we are far from exploiting the real power of those machines. However, they really helped us to do our research effectively.

References:

[1] Than K., Ho TB., “Modeling the diversity and log-normality in data”, Submitted

5.　2009 年度の計算サーバ利用研究

A study on RC4 key collision

Jiageng Chen, Atsuko Miyaji

Jun 30, 2010

Stream ciphers is a kind of symmetric key ciphers where users use the same key for encryption and decryption. Plaintext message bits (bytes) are encrypted (decrypted) one at a time by combining with a pseudorandom cipher bit stream (keystream). Our research focuses on cipher RC4 which is one of the most famous world wide used stream ciphers. It is used in real world applications such as Microsoft Office, Secure Socket Layer (SSL), Wired Equivalent Privacy (WEP), etc. RC4 has become a hot cryptanalysis target since its specification was made public on the Internet in 1994. Various general weaknesses of RC4 have been discovered in some previous works. Another popular cryptanalysis direction of RC4 is in the WEP environment.

Our research mainly focuses on RC4 key collisions. It is about the existence of secret key pairs that generate the same initial states after key scheduling algorithm. Thus from the cryptographic point of view, it is totally indistinguishable. This is a serious weak point for a stream cipher and thus should be studied carefully. The study of "colliding keys" of RC4 can be dated back to 2000. Grosul and Wallach first pointed out that RC4 exists near collisions (two internal states are almost the same with some few differences) when the key size is close to the full 256 bytes. In FSE2009, Matsui first found the total colliding key pairs with hamming distance one (Two keys differ from each other at one position).

Our main contribution this year is the discovering of a new class of colliding key pairs of RC4, namely, two different secret keys generate the same internal state after RC4's key scheduling algorithm. This is to our knowledge the first discovery of RC4 colliding keys with hamming distance greater than one, that is, the colliding key pairs we found can differ from each other at three different positions, and the value difference between these positions needs not be fixed. We analyzed the transition pattern and evaluated the

probability of the existence of this new class of colliding key pairs. Our result shows that RC4 key collision could be achieved by two keys with greater hamming distance than the ones found in Matsui's paper in FSE2009. And this newly discovered class of colliding key pairs reveals the weakness that RC4's key scheduling algorithm could generate even more colliding keys. We also proposed an algorithm for searching colliding key pairs within this new class and by using our algorithm, a 55-byte colliding key pair is the shortest one we found within one hour time.

- Jiageng Chen and Atsuko Miyaji. "A New Class of RC4 Colliding Key Pair",the 2010 Symposium on Cryptography and Information Security, SCIS2010, 3D2-4.

- Jiageng Chen and Atsuko Miyaji, "A New Class of RC4 Colliding Key Pairs With Greater Hamming Distance", The 6th Information Security Practice and Experience Conference, ISPEC 2010, Lecture Notes in Computer Science, 6047 , Springer-Verlag, 2010.

6. 謝辞

本報告は、JAIST の並列計算機利用者メーリングリストである mpc メーリングリストにおいて各ユーザへ報告を依頼し、各著者のご厚意によって提出いただいた文章をまとめたものである。ご多忙のところ、ご協力をいただいた著者各位に感謝します。