

CMSI計算科学技術特論C(第8回)2015年11月19日

# アウトソーシングによる シミュレータの共同開発

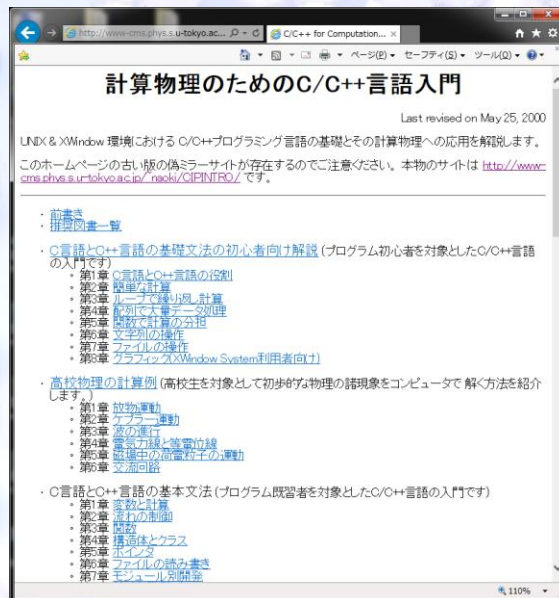
みずほ情報総研株式会社  
サイエンスソリューション部  
渡辺 尚貴

---

# 自己紹介(学部学生の頃)

1994年～1998年 東京大学 理学部 物理学科

趣味としてプログラム技術を学び、ホームページにまとめて公開。



インターネット創生期のHP

C++言語入門、計算物理、X11グラフィックス  
TCP/IP通信、CGI Webプログラム、LaTeXなど

自分が書いた文章・プログラムが社会に貢献することを実感。

2015年現在も残る。 <http://www-cms.phys.s.u-tokyo.ac.jp/~naoki/>

# 自己紹介(大学院生の頃)

1998年～2003年 東京大学大学院 理学系研究科 物理学専攻 塚田研究室

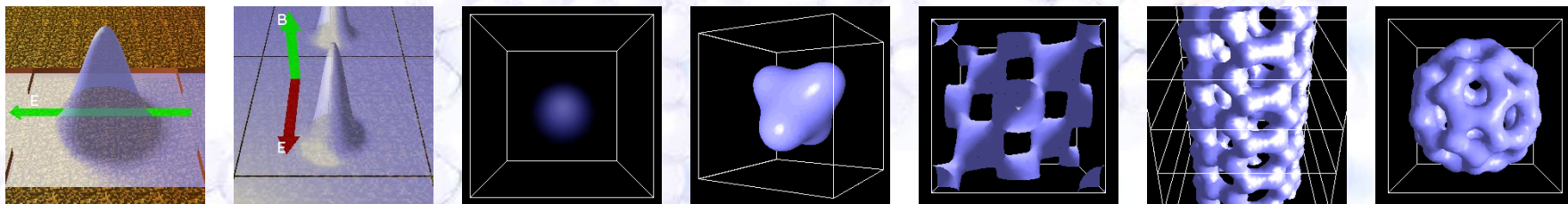
時間依存Kohn-Sham方程式の第一原理計算を研究(学振DC1)。

$$i\hbar \frac{\partial}{\partial t} \phi_n(\mathbf{r}, t) = \left[ -\frac{\hbar^2}{2m} \Delta + V_{\text{ions}}^{\text{PPloc}} + \hat{V}_{\text{ions}}^{\text{PPnon}} + V_{\text{H}} + V_{\text{XC}}[\rho] + V_{\text{ext}}(t) \right] \phi_n(\mathbf{r}, t) \quad \rho(\mathbf{r}, t) = \sum_n f_n |\phi_n(\mathbf{r}, t)|^2$$

実空間・実時間法での高精度・高安定な計算アルゴリズムを考案。

$$\begin{aligned} \phi_n(\mathbf{r}, t + \Delta t) = & \exp\left[i\frac{\Delta t}{4} \Delta\right] \exp\left[-i\frac{\Delta t}{2} V_{\text{ions}}^{\text{PPloc}}\right] \exp\left[-i\frac{\Delta t}{2} \hat{V}_{\text{ions}}^{\text{PPnon}}\right] \exp\left[-i\Delta t(V_{\text{H}} + V_{\text{XC}}[\rho] + V_{\text{ext}}(t))\right] \\ & \times \exp\left[-i\frac{\Delta t}{2} \hat{V}_{\text{ions}}^{\text{PPnon}}\right] \exp\left[-i\frac{\Delta t}{2} V_{\text{ions}}^{\text{PPloc}}\right] \exp\left[i\frac{\Delta t}{4} \Delta\right] \phi_n(\mathbf{r}, t) \end{aligned}$$

シミュレータを開発して、簡単な実証例を示して卒業(理学博士)。



趣味として時間発展する波動関数の可視化プログラムも作成。



# 自己紹介(会社での業務)

2003年～ みずほ情報総研株式会社 サイエンスソリューション部

- 量子電子動力学シミュレータの開発・計算
- 電子状態計算プログラムの開発・改良・並列化
- 電子状態計算プログラムの説明書整備
- 電子状態計算プログラムのフロントエンド開発
- 分子動力学計算アルゴリズムの改良
- タンパク質構造解析プログラムの改良
- 燃料電池シミュレータの開発
- 走査プローブ顕微鏡シミュレータの開発
- ナノテクロードマップの作成

# 会社紹介 みずほ情報総研株式会社

## 主力業務

みずほ銀行の銀行システム開発。 約4000人

## その他業務

企業、官公庁のコンサルティング。 約500人

分野： 経営相談、環境評価、社会政策、情報通信、**科学技術**

## 科学技術コンサルティング業務(サイエンスソリューション部) 約60人

大学・国研の研究、企業研究所の開発、官公庁の政策をIT技術で支援。

分野： 原子力、防災、土木、自動車、電池、半導体、医薬、**材料(4人)**

# 部署紹介 サイエンスソリューション部

## 主力業務

大学・国研・民間の研究者・技術者への科学コンサルティング。

## 研究課題

顧客の研究課題の解決を支援。

課題解決のための技術と知見の蓄積が豊富。

## ソフトウェア

新規ソフトウェアを自社で開発して顧客に納品。

市販ソフトウェアを使用して計算結果を顧客に報告。

市販ソフトウェアとサポートを併せて顧客に販売。

## スタッフ

数学、物理、化学、生物、工学の修士・博士卒の技術者約60人



# 科学技術コンサルティング

## 1. 科学技術分野の動向調査

官公庁、メーカーの依頼で、施策決定材料とする。  
技術動向・予測資料を研究者ヒアリングから作成。

## 2. シミュレータを用いた解析

大学・国研の実験研究者、メーカーの技術者の依頼で、  
顧客課題の現象を既製品でシミュレーション解析し、結果を報告。

## 3. シミュレータの開発・改良(本講義での主題)

大学・国研の理論・計算研究者、メーカーの技術者の依頼で、  
顧客アイデアをプログラム化、顧客開発のプログラムを改良して納品。

# シミュレータ開発改良の依頼主の事情

(大学・国研の研究者の事情)

- プログラム開発が好きなので自分で開発して研究したい。
  - 多忙(教育、公募、学科、学会、雑用)のため開発の時間は不足。
  - ポスドク・学生に開発させたいが研究課題には向かない。
  - 学生がプログラミング言語を修得するのは容易ではない。
  - 学生に開発させたプログラムの開発ノウハウが学生が卒業して消失。
  - 新しいプログラミング技術を追いかける時間が不足。
  - 単独で開発してきたが、長年の複雑化で開発が行き詰まり。
  - プログラム開発者としての評価は論文著者としての評価より低い。
  - 公募の研究予算を獲得した。計算機を買うか、ポスドクを雇うか迷う。
- 上記理由によりプログラム開発の業者(当社)への外注を考え始める。



# 理論計算研究者の研究工程

1. 解明したい現象の選定
2. 現象の解明方法の策定
3. 基礎方程式の策定
4. 計算アルゴリズムの選定
5. 計算プログラム言語の選定
6. データ構造の設計
7. データ処理の設計
8. プログラムの実装
9. プログラムのデバッグ
10. プログラムの並列化・高速化
11. プログラムの説明書作成
12. プログラムのテスト系での計算
13. 計算結果のプロット・可視化
14. プログラムの本番系での計算
15. 計算結果の物理的価値の発見
16. 学会発表・論文投稿

研究者自ら担わずとも外注業者が分担できる研究工程も存在する。

# 研究者と当社との分担割合

- |                 |                    |
|-----------------|--------------------|
| 1. 解明したい現象の選定   | 9. プログラムのデバッグ      |
| 2. 現象の解明方法の策定   | 10. プログラムの並列化・高速化  |
| 3. 基礎方程式の策定     | 11. プログラムの説明書作成    |
| 4. 計算アルゴリズムの選定  | 12. プログラムのテスト系での計算 |
| 5. 計算プログラム言語の選定 | 13. 計算結果のプロット・可視化  |
| 6. データ構造の設計     | 14. プログラムの本番系での計算  |
| 7. データ処理の設計     | 15. 計算結果の物理的価値の発見  |
| 8. プログラムの実装     | 16. 学会発表・論文投稿      |

各研究工程の分担割合の模式図(■研究者の分担割合 ■当社の分担割合)

研究者は研究活動に専念、当社は開発に専念。相互協力で補完。

部分的な外注で研究開発が進展。

# 外注業者が開発できる理由

## 当社コンサルタントの出身分野

数学、物理、化学、生物、工学の分野の大学院。修士・博士号取得卒。

## 当社コンサルタントの開発経験

構造、流体、半導体、バイオ、物性(第一原理計算)のプログラム開発。

どの分野のプログラムも配列データのループでの計算処理に帰着。

## 最先端の科学に対するプログラム開発

最先端の理論や計算式は研究者から提供を受ける。

後は計算式に従い行列・ベクトルの反復計算。

大規模系では配列サイズとループ回数が多くなる。

計算内容は理解できる。その理論的背景もある程度理解できる。



# コンサルティング業務の流れ(1)

## 顧客から問い合わせ(HP、メール、電話)

依頼内容の概要を受け、適任担当者を決めて返答する。

## 顧客先を訪問

研究課題の説明を受け、依頼内容の詳細を確認する。

## 対応可否を検討

依頼内容を引き受けられるか社内で検討する。

## 契約書類を作成

依頼内容を仕様書として文章化。

作業工数(人数、時間、納期)を予想し、見積書を決定。

顧客に提示し、了解を得る。

## 契約を締結

顧客の所属機関の発注部門と当社の契約担当間で契約手続きする。

# コンサルティング業務の流れ(2)

## 業務の開始

開発に必要なデータ(コード、マニュアル)を受領する。

社内計算機を用いて開発する。

## 定期的な進捗報告

1～2カ月おきに顧客先を訪問し、順調な進捗を報告する。

仕様の実装法の細部を協議して決定する。

## 納期での納品

納品物(コード、マニュアル、請求書)を持参して顧客先を訪問し、

納品物の内容・取り扱い方法について説明する。

## 納品後の対応

バグを修理し、質問へ回答する。

# 業務例 新規プログラムの開発

**顧客:** 物質・材料研究機構(NIMS)

**年度:** 2006年度～

**目的:** NIMS公開中の量子伝導度計算プログラムASCOT用のハミルトニアン行列要素を計算するプログラムACCELの開発

**作業:** 顧客側から電子状態計算に関する論文が提供。  
その論文の機能をFortran90プログラムとして実装。  
1カ月おきに進捗報告打合せ。次の機能に関する論文が提供。  
この繰り返しで数年かけて完成。現在NIMSから公開中。

**機能:** 行列要素、セルフコンシステント場、固有エネルギー、固有波動関数、バンド図、状態密度、中性原子ポテンシャル、周期遮蔽ポテンシャル、原子間力、構造最適化、スピン軌道相互作用、MPI並列化(ScaLapack使用)

**納品物:** ソースコード一式、理論説明書、操作説明書、プログラム設計書



# 業務例 顧客プログラムの並列化

**顧客:** さまざまな分野の研究者

**作業:** 顧客からコードとマニュアルを受領。

コードの計算処理をサブルーチン、モジュールを追いかけて分析。

コードの計算処理を並列化しやすいように整理。

並列可能なループを特定し、OpenMP並列化指示行を記述。

分散可能なデータを特定し、MPI通信関数を記述。

並列計算の動作検証。

並列計算の性能検証。

**納品物:**

ソースコード一式、

作業報告書

**その他並列化:**

ハイブリッド並列化、

高次並列化、GPGPU、XeonPhi

```
real(8) :: data( mpi%this_start, mpi%this_end )
```

```
!OMP PARALLEL DO PRIVATE(i)
```

```
do i=mpi%this_start, mpi%this_end
```

```
    data(i) = ...
```

```
end do ! i
```

```
!OMP END PARALLEL DO
```

```
call MPI_Allgatherv( data, mpi%this_size, ..., &  
    data_all, mpi%each_size, mpi%each_position, ... )
```

# 業務例 入力ファイル形式の整備

## XML形式の入力ファイルの提案

```
<lattice>  
  <unitvector unit="angstrom"> 5.43 0.00 0.00 </unitcell>  
  <unitvector unit="angstrom"> 0.00 5.43 0.00 </unitcell>  
  <unitvector unit="angstrom"> 0.00 0.00 5.43 </unitcell>  
</lattice>
```

## XML形式の利点

- 人間の目で読みやすい(量が少ないうちは)。
- プログラムでも読みやすい(F90用のライブラリ(xmlf90)もある)。
- 人間の手で書き込みやすい(雛形の値を変えるだけなら)。
- XMLを編集するフリーソフト(XML Notepad)もある。
- 量が多い入力データ(原子座標、行列要素など)は別ファイルにすれば良い。

## XML化の作業

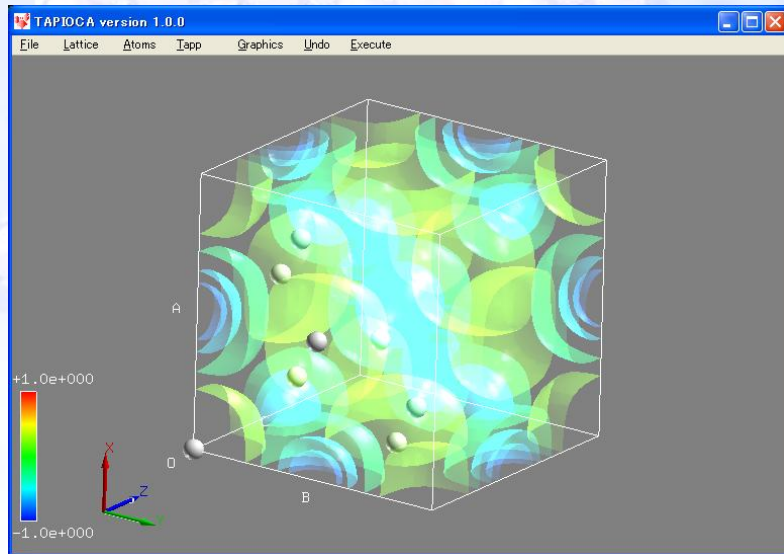
パラメタの策定。

読み込みプログラムの作成(xmlf90改変版を使用)。

```
vnode => getElementByTagName(node_lattice,"unitcell")  
call getChildValueChar( value, item(vnode,0) )  
read(unit=value,fmt=*) lattice%Ax, lattice%Ay, lattice%Az
```

# 業務例 専用可視化ソフトウェアの開発

顧客のプログラムの計算結果を描く可視化ソフトウェアを多数開発。



東京大学常行研究室に納品した  
xTAPPフロントエンド TAPIOCA の  
グラフィックス画面

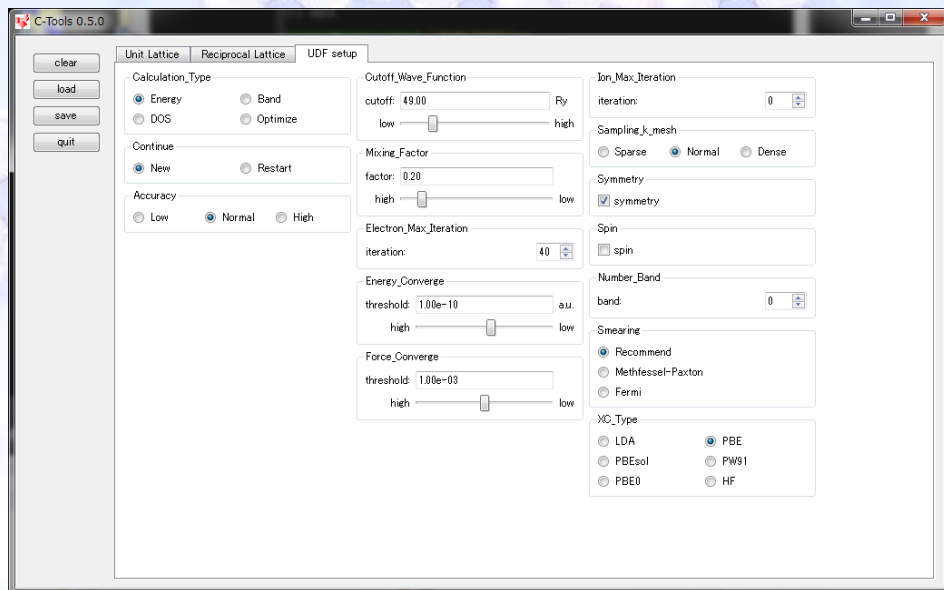
## OpenGLライブラリ

三次元グラフィックスソフトウェアを開発するための業界標準ライブラリ。  
単純な多角形(ポリゴン)の描画機能がある。  
等数値面などを描くには複雑なプログラミングが必要。  
計算結果の空間分布量などを瞬時に視覚的に理解できる。



# 業務例 専用GUIソフトウェアの開発

顧客のプログラムの計算内容を設定するGUIソフトウェアを多数開発。



東京大学常行研究室に納品した  
第一原理計算変換アプリC-toolsの  
GUI画面

## Qtライブラリ

GUIソフトウェアを開発するためのライブラリ。

選択ボタンや数値入力欄などのさまざまなGUI部品がある。

GUI部品と計算パラメタを関連付けるには複雑なプログラミングが必要。

計算内容をマウスだけで簡単に設定することもできる。

# プログラム納品のその後

研究者がソースコードを活用して研究を加速。

研究者が計算結果から物理的価値を見出し、論文を量産。

他の研究者から注目、利用者が増加、引用数が増加。

製造業の技術者から注目、利用され、新製品開発・新産業にも貢献。

しかし、計算できるだけではプログラムは普及しない。

プログラムが普及しない原因がある。

プログラムが普及しない原因を除く別の取り組みが必要。

# プログラムが普及しない原因

製造業の技術者もそのプログラムを利用したいが入手できない。

技術者が抱える課題にも適応できるか技術者にはわからない。

ニッチな機能が必要な課題である。高精度でなくても良い。

日本製の無償のプログラムを使用することに不安がある。

コンパイルでエラーになる。実行するとエラーになる。

入力ファイルのパラメタ指定方法がわからない。

計算結果の評価方法がわからない。

研究者にノウハウを質問したいが、計算内容は公開できない。

大型計算機を利用できない。計算結果は公開できない。使い方がわからない。



# プログラムの普及への取り組み

製造業の技術者もそのプログラムを利用したいが入手できない。

**配布ホームページを開設する。配布法人を作る。**

技術者が抱える課題にも適応できるか技術者にはわからない。

**解析事例集を作成する。**

ニッチな機能が必要な課題である。高精度でなくても良い。

**個別のニーズに応える追加機能を開発する。**

日本製の無償のプログラムを使用することに不安がある。

**論文で実績を示す。企業での仕様実績を示す。**

コンパイルでエラーになる。実行するとエラーになる。

**Makefile等を整備する。標準文法に従うコードにする。各環境でテストする。**

入力ファイルのパラメタ指定方法が分からない。

**操作説明書を作成する。GUIを作る。自動設定にする。**

計算結果の評価方法が分からない。

**可視化の仕組みを作る。評価方法の相談に応じる。**

研究者にノウハウを質問したいが、計算内容は公開できない。

**機密保持契約のうえ利用できる相談窓口を設ける。**

大型計算機を利用できない。計算結果は公開できない。使い方がわからない。

**企業でも利用できて公開不要な大型計算機もある。講習会を開く。**

# まとめ

外注業者でも研究用プログラムを開発できる。

開発を業者に外注すれば、研究者は研究に専念できる。

物理的価値のある計算成果が早く得られ、研究者は研究を加速できる。

プログラム開発を専門家に託すことで、プログラムがより発展する。

プログラムを普及させるには、そのための取り組みが必要である。

プログラムの利用者が増えれば、論文引用数が増加する。

プログラムが産業に利用されれば、社会貢献となる。

ご清聴ありがとうございました。

当社科学技術コンサルティング紹介 & お問い合わせ窓口

<http://www.mizuho-ir.co.jp/solution/research>