

グローバルコンピューティング環境による汎用MCSCFソルバーの開発

日本電気(株)システムデバイス・基礎研究本部 高田 俊和

Development of an MCSCF Program System Adjusted to Global Network Computing Environments

Toshikazu Takada, System Devices and Fundamental Research, NEC Corporation

Since chemical reactions are widely used in productions of industries as is seen in semi-conductor device processes, it is needed to develop molecular simulation programs which are able to analyze the chemical reaction mechanisms, based on quantum mechanics. In this project, therefore, an MCSCF (Multi Configuration Self Consistent Field) program including energy gradient calculations is newly developed, by using parallel computing algorithms. Furthermore, to make it easy to use the MCSCF program system, a GUI is developed as well, by which researchers make input data, carry out calculations and analyze results through Internet. This system will propose future images of molecular simulation.

1. はじめに

産業における製造プロセスがプレスなどの物理的加工から、半導体製造の微細化に見られるように、化学反応を利用した化学的加工へ工業技術の高度化と共に移行しつつある。従って、このような化学反応のメカニズムを原子・分子レベルで解明できるシミュレーションシステムを開発することは、新素材の開発や生産工程の効率化などに極めて有効であると期待されている。他方、タンパク質の化学反応を人工的に改変し、産業的に活用することが現実味を帯びてきているが、これらバイオ関連分野においても分子シミュレーションの必要性が高まっている。

しかしながら、量子力学に則り化学反応をシミュレートしようとする、非経験的分子軌道法として確立されている計算手法の中でも、MCSCF (Multi Configuration Self Consistent Field)法による計算が不可欠である。その理由は、化学反応の進行と共に変化する電子配置の混合を記述できるからである。よって、化学反応の機構解明に必要な活性化エネルギーの計算が正確に行なえるので、溶液中や生体分子の活性部位での複雑な化学反応のメカニズムを解明するためのコアテクノロジーとして位置づけられる。しかしながら、既存のMCSCFプログラムの適用範囲は、アルゴリズム的な制約から小型分子に限られる場合が多く、新素材開発を行なっている研究現場での要請に応えられる段階には到達していない。

コンピュータのハードウェア性能の向上は、並列化をベースに進められている。また、インターネットの著しい普及に伴い、シミュレーションを含めコンピュータの利用形態が大きく変化している。これからのシミュレーションはインターネットを介して行なわれると考えられるので、これらのハードウェア動向に着目しながら、本研究課題では、1) CASCF (Complete Active Space SCF)法に基づいて、生体分子などの大型分子のMCSCF計算を高速通信回線で接続されたヘテロジニアスなスーパーコンピュータで実行する数値計算部の開発、2) インターネットを介して計算結果を複数の研究者が各自のパソコン上に表示し、且つ会話を通して同時に解析できるステアリングシステムの開発、の2点を中心テーマとして行なう。このようなシステムはまだ報告されておらず、分子シミュレーションシステムの将来的イメージを世界に先駆けて提案できると考えている。

2. MCSCF理論の概略とプログラム開発

非経験的分子軌道法の理論について、その概略を述べる。量子力学に則り、シュレーディンガーの波動関数は、

$$(1) \quad H \psi = E \psi$$

$$(2) \quad H = -\frac{1}{2} \nabla^2 - \sum_a \frac{Z_a}{r_{ia}} + \sum_{i,j} \frac{1}{r_{ij}} + \sum_{a,b} \frac{Z_a Z_b}{R_{ab}}$$

$$(3) \quad \psi = \sum_i C_i \phi_i$$

で与えられる。ここで、HはHamiltonian演算子、 ψ 及びEは解として得られる波動関数及び全エネルギーである。

はパウリの原理を満たすように反対称化されたスレーター行列式で、

$$(4) \quad \Phi = 1/\sqrt{N!} |\dots \phi_{i/2}(i-1) \alpha(i-1) \phi_{i/2}(i) \beta(i) \dots|$$

のように与えられる。1電子関数(分子軌道)は、既知の原子軌道の線形結合で

$$(5) \quad \psi_i = \sum_r C_{ri} \phi_r$$

と定義される。ここで導入された2種類の係数C、cの微小変化に対して、期待値

$$(6) \quad E = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle}$$

が停留値をとるという条件から、係数を求める方程式を導出することができる。これが、MCSCF法の基本的な考え方である。

一方、原子核に働く力は、この全エネルギーEを原子核の座標aで微分して得られるので、その表式は、

$$(7) \quad \frac{\partial E}{\partial a} = \sum_{rs} \gamma_{rs} \frac{\partial h_{rs}}{\partial a} + \sum_{rstu} \Gamma_{rstu} \frac{\partial (rs|tu)}{\partial a} - \sum_{rs} W_{rs} \frac{\partial S_{rs}}{\partial a}$$

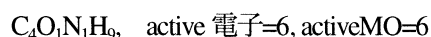
で与えられる。演算の大部分が2電子積分と各座標による微分の計算に費やされており、大型分子の計算を実現するには、それらの積分計算を如何に並列化し且つインコアで処理するかが重要な課題となる。本プロジェクトでは、この並列化に十分配慮してプログラム開発を進めており、既存のプログラムでは計算できない大型分子のMCSCF計算が可能になると考えている。

3. 開発経過

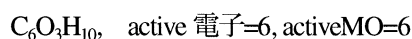
MCSCF法における計算プロセスでは、電子配置の係数Cを求める配置間相互作用計算と分子軌道の係数cを求めるSCF計算を、これらの係数が前回求めた値と一致するまで繰り返す。この数値計算部をデバッグするため、既存のプログラムGAMESS(米国製)を用いて全エネルギーを比較したところ、

分子	GAMESS	本プロジェクト版	(原子単位)
H ₂ O	-75.9615241386	-75.9615241386	
H ₂ CO	-113.8014219135	-113.8014219135	
NaOH ₂	-237.8331721124	-237.8331721124	

となり、CI行列要素及び一般化Fock行列の計算など電子配置を扱うコアになる部分については正しく動作していることが確認された。収斂性能をNEC製スーパーコンピュータSX-4/1で測定したところ、下記の表のようになっている。収斂回数も少なくベクトル性能も出ているので、プログラムのコーディング自体は概ねうまくいっていると考えている。しかしながら、2電子積分の取り扱いを従来のアルゴリズムで単純にコーディングすると、表に示されているように主記憶上の作業領域不足から計算できない場合が起こる。それは、2電子積分の数が用いる原子軌道(AO)の4乗に比例して増加するという、非経験的分子軌道法の本質的な問題に由来している。本プロジェクトでは、2電子積分駆動型アルゴリズムの採用と並列分散処理によるローカルメモリの活用とにより、既存のプログラムでは計算できない1000軌道以上の計算を実現することを目的としている。



	AO 数	SCF 回数	User Time(sec)	MFLOPS	メモリ(MB)
MINI-4	39	12	157	305	500
MIDI-4	72	20	712	484	500
MIDI-4**	135	20	4960	602	2000



	AO 数	SCF 回数	User Time(sec)	MFLOPS	メモリ(MB)
MINI-4	55	11	233	699	500
MIDI-4	101	18	2063	751	2000
MIDI-4**	185		メモリー不足		2000

エネルギー勾配計算部については、CASSCF法によって与えられるエネルギー表式を原子核の座標で微分した式を求め、計算手順を確認した。現在、この表式に基づいて、プログラムのコーディングを進めている。設計面で工夫した点は、原子軌道基底で計算される2電子積分の核座標による微分を分子軌道基底に変換する従来のアルゴリズムとは異なり、分子軌道基底で与えられる表式を逆変換して、積分値との積和をとるようにしたことである。これにより、計算時間の短縮と主記憶上での作業領域の削減ができると考えている。

4. ネットワークの活用

シミュレーション結果の解析には、分子グラフィックスは必須である。特に大型分子の計算では分子軌道の係数からその形状を想像するのは不可能に近く、最近の論文では分子軌道を分子グラフィックスで表示している例が多い。現行の分子グラフィックスの多くが、高速CGエンジンを搭載しているグラフィックスワークステーションにグリッドデータを送り、回転や拡大をローカルな処理として行なっている。しかしながら、この方法では転送すべきデータ量が多く、スーパーコンピュータとグラフィックス端末が近接していなければならないという制約がある。分子シミュレーションの多くが今後インターネットを介して行われるようになると考えられるが、グリッドデータの転送は負荷が大きく効率的でない。数値計算を実行するコンピュータ上でピクセルデータにまで変換してしまえば、画像圧縮技術により転送データ量を著しく削減でき、インターネットによる表示が十分可能となる。回転などのインタラクティブな操作は、手許のパソコンからピクセルデータを作成するコンピュータに指示を出せば良いことになる。

異なる分野の研究者が、同一の分子の性質を色々な観点から研究することは良くある。計算結果に関する情報をこれらの研究者間で効率良くシェアできるようにするため、複数のパソコンから同時にブラウザを起動し、同一画面を見ながら議論できるGUIが整備されれば、遠く離れた研究者間の共同研究を促進できると期待される。以上の点を考慮し、多地点の研究者がインターネットを介してMCSCFの計算結果をグラフィックス画面と会話とにより解析できるGUIシステムを開発している。図1は、GUIのブラウザ画面である。現在、複数拠点からアクセスし、動作確認を進めているところである。



図1 GUIのブラウザ画面

5. 今後の予定

数値計算部については、次のように作業を進める予定である。まず、波動関数計算部の並列化について、新しく考案したアルゴリズムによるプログラムのデバッグ作業とベンチマークを行い、その実効性能を確認する。また、1000軌道以上の計算能力を示すための具体的な応用例としてクロロフィル2量体の励起状態の計算を現在検討しており、プログラムのテストが終了し次第計算を始めたいと考えている。エネルギー勾配計算部については、現在プログラムのコーディングを進めており、今年度中にはデバッグ作業を終了する予定である。更に、中程度の大きさの分子でのベンチマークを行い、並列性能などを確認する。一方、GUIと数値計算部のインターフェースをとり、開発されたプログラムを実際に協力研究員にインターネットを介して利用して頂き、使い勝手などについても評価し改良を行なうことを考えている。

6. 研究実施体制

研究開発題目：汎用MCSCFソルバー及びインターネットGUIの開発

研究開発題目：汎用MCSCFソルバー及びインターネットGUIの開発

氏名	所属	役職	研究開発項目
高田 俊和	日本電気基礎研究所	主席研究員	MCSCF計算部設計
津田 健一郎	日本電気基礎研究所	研究員	分子グラフィックス

研究開発題目：MCSCF分散処理のための基本アルゴリズム開発及びIMnetへの接続によるプログラム評価

氏名	所属	役職	研究開発項目
関口 智嗣	工技院電総研	主任研究官	分散処理アルゴリズム
建部 修見	工技院電総研	研究員	分散処理アルゴリズム

研究開発題目：IMnetへの接続によるプログラム評価

氏名	所属	役職	研究開発項目
柏木 浩	九工大情報工学部	教授	IMnetによる性能評価

研究開発題目：IMnetへの接続によるプログラム評価

氏名	所属	役職	研究開発項目
日向寺 祥子	東海大学計算センター	助手	IMnetによる性能評価