

# 並列分子動力学法シミュレーションの 動的負荷分散法

林 亮子      堀口 進

北陸先端科学技術大学院大学 情報科学研究科

物理、化学の分野で用いられる数値シミュレーションの一手法である分子動力学法の並列計算に関する研究は、現在盛んに行われている。Coarse-grained cell method は、相互作用の局所性を用いて並列計算を行なう効率的な方法として期待されている。しかし、シミュレーションの物理条件によって、特定のプロセッサに計算負荷が集中し、並列計算の効率が極端に悪くなる。本研究では Coarse-grained cell method を用いた大規模分子動力学法シミュレーションにおいて、セルを単位としたタスクスケジューリング法による動的負荷分散法を提案する。また、動的負荷分散法を超並列コンピュータ CM-5 上に実装し、その有効性について詳しく検討する。

## Dynamic load balancing for parallel molecular dynamics simulation

Ryoko Hayashi      Susumu Horiguchi

School of Information Science  
Japan Advanced Institute of Science and Technology

Studies on parallel computing have become very important for molecular dynamics simulation in physics and chemistry. Coarse-grained cell (CGC) method is expected as a powerful parallel simulation strategy, because CGC uses a locality of mutual action between molecules. However, computational load unbalance on specific processing elements makes the efficiency of parallel computation worse. We propose a method of dynamic load balancing for CGC parallel simulation by introducing a task scheduling strategy. The performance of dynamic load balancing are discussed by implementing the proposal strategy on a massively parallel machine CM-5.

## 1 はじめに

近年、コンピュータの高速化にともなって先端科学技術分野でのコンピュータ・シミュレーションの果たす役割が大きくなっている。ここでは、そのような数値計算の手法の一つ、分子動力学法 [1] の並列シミュレーションについて研究結果を報告する。

分子動力学法で扱われている分子の数は逐次処理で数千個から、非常に簡単な相互作用の場合で百万個程度である。本来物質はアボガドロ数  $6.02 \times 10^{23}$  個程度の分子の集まりである。それに比較して現在計算可能な物質は量的、質的に限られたものであり、並列コンピュータによる高速化が期待されている。

これまでに PE 数、CPU、結合方式の面で様々な並列計算機を用いた研究が行なわれている。並列化に際して、空間を分割してプロセッシングエレメント (PE) に割りあてる方法と、位置によらず分子の数を分ける方法が比較されている [2]。Beazley, Lomdahl [3] は 1024PE の超並列計算機上で、空間を分割して PE に割り当てる方法 Coarse-grained cell method (以下 CGC 法と称す) を用いて  $6.7 \times 10^7$  分子を計算している。Hedman [4]、Tamayo [5] 等は、CGC 法をベースに、さらに計算量を減らす手法を検討している。また CGC 法とは異なるが、空間を分割して PE に割りあてて動的負荷分散を考慮し、2次元空間のシミュレーションが行なわれている [6]。

しかし、これらの研究では、シミュレーションの物理条件によって処理効率に変化する問題について十分に研究されていない。CGC 法では、低温低密度の物理条件で高速化の効果が減少する。本論文では、短距離相互作用の分子動力学法シミュレーションにおける動的負荷分散法を提案する。さらに超並列コンピュータ CM-5 [7] 上に実装し、その有効性について議論する。

本論文の構成は以下の通りである。第 2 章では分子動力学法に簡単に触れる。第 3 章では CGC 法について述べ、動的負荷分散法を提案する。第 4 章では提案した動的負荷分散法を並列計算機 CM-5 上に実装し、その有効性について議論する。第 5 章はまとめである。

## 2 分子動力学法

分子動力学法とは、物質を構成する分子一つ一つの運動を模擬することによって、物質の状態を調べる計算手法である。運動方程式としては多くの場合、

ニュートンの運動方程式を用い、分子に働く力は分子の 2 体力の和によって与える。そのとき扱う分子の数を  $N$  とすると、力の組みあわせは

$${}_N C_2 = \frac{N(N-1)}{2}$$

通りである。そのため、一般に計算時間は  $O(N^2)$  で増加する。ここでは分子の間に働く力は、ポテンシャル関数で表した分子間相互作用として与えられる。分子間距離が増大すると分子間相互作用の大きさが急速に 0 に近づく場合には、ある距離以上では相互作用を零として、力の計算時間を短縮する。このとき、相互作用の有効範囲を「力の切断距離」という。このような相互作用は短距離相互作用と呼ばれる。本研究では短距離相互作用を扱う。

ここでは、相互作用としてレナード・ジョーンズ型ポテンシャル

$$V(r) = 4\epsilon \left\{ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right\}$$

を用いる。ただし、 $\epsilon$ 、 $\sigma$  は物質によって異なる定数である。このポテンシャルはアルゴンやヘリウムなどの、電気的に中性な稀ガス元素の単原子分子の気体および液体について、適用されてきた球対称ポテンシャルである。ここではアルゴンの値を使用し、物理条件としては、分子の数  $N$ 、体積  $V$ 、温度  $T$  を一定とした。分子の直径を単位として、力の切断距離は通常 2.5 ~ 3.5 が用いられるが、ここでは 2.5 を使用した。

時刻  $t$  での位置  $r(t)$ 、速度  $v(t)$  の時間発展に用いる数値積分は、ベルレの方法の速度形式 [1] を用いた。

## 3 分子動力学法の並列シミュレーション

### 3.1 並列計算機環境

本研究では 64PE の並列計算機 CM-5 上に実装を行なった。CM-5 [7] は PE に SPARC microprocessor (浮動小数点演算ユニット、33MHz, 22MIPS, 4.2MFLOPS) を使用し、1PE に 4 基のベクトルユニット (4 基で 16MHz, 32MFLOPS) を持つ。結合網は fat tree であり、各 PE は tree の葉にあたる。そのため、PE 間通信において、PE が通信を中継する必要はない。ここでは FORTRAN の並列拡張言語 CM Fortran に CMMD ライブラリを併用し、MIMD 型処理を行なった。

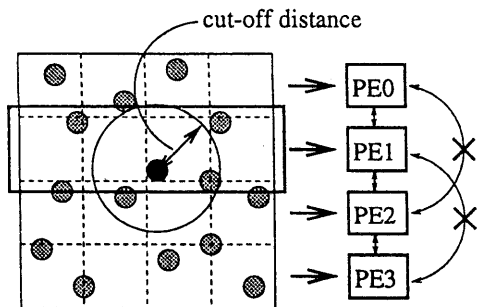


図 1: Coarse-grained cell method(CGC 法)の概要

### 3.2 並列化手法

CGC 法は空間を立方体型に分割 (セル) し、セル単位で PE に割り当てる方法である。2 次元空間のシミュレーションについて、CGC 法の概要を図 1 に示す。力の切断距離よりもセルが大きければ、各セル中の分子は近傍セル (2 次元空間では 8 近傍、3 次元空間では 26 近傍) 中の分子との相互作用を考慮すればよい。本研究では CGC 法を用いるが、特に各 PE の持つセルは空間的にまとまって柱状をなすものとする。各 PE は 8 近傍 PE とのみ局所通信を行なう。

低温、低密度の状態のシミュレーションを行なう場合、CGC 法には不都合が生じる。分子が空間の一部分に集中すると、セル中の分子の数にばらつきがあるので、PE ごとの計算負荷にも不均衡が生じる。シミュレーションで現れる、分子が集中した状態の例を図 2 に示す。物理量を計算するために、分子動力学法では各時間刻みごとに全 PE で同期をとる。そのため、PE に計算負荷の不均衡が生ずると、計算効率は悪化する [3]。分子の集中度および集中する場所は刻々変化し、予測不可能である。したがって、CGC 法では動的に負荷分散を行なう必要がある。

### 3.3 動的負荷分散法

空間を PE に割り当てる手法における動的負荷分散法は、2 次元のシミュレーションについて検討されている。2 次元空間を短冊状に分割して PE に割り当て、分子の数を均等にするように PE の担当する空間の境界を移動させる方法が報告されている [6]。一方 CGC 法では、セルの境界が変化するとセルの空間的配置が変わり、26 近傍のみを参照する局所条件が成り立たなくなる。したがって、分割した空間

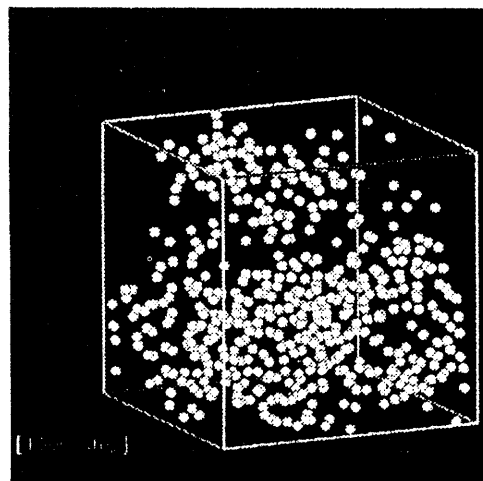


図 2: 低温、低密度で現れる分子の状態

の境界が変化する方法を単純に拡張することはできない。

本研究では、大規模シミュレーションにおいてセル数が膨大になることに着目し、セルを単位としたタスクスケジューリングによる動的負荷分散法を提案する。この手法では、計算負荷の高い PE は近傍 PE に分子のデータをセルごと送って (セルの再分配) 全体として計算負荷の不均衡を小さくする。

セルの再分配の概要を図 3 に示す。初期状態では、全ての PE は同じ数のセルを同じ形にまとめて持つ。セルの再分配の結果、図 3(b) に示したような場合は PE の隣接関係が保持されて、各 PE は 8 近傍 PE のみと通信すれば良い。一方図 3(c) のように再分配された場合は、近傍 PE 数が変化し、8 近傍 PE のみと通信する局所条件が成立しない。ここで、図 3(d) に示すように PE 固有のセル (固定分配セル) を定義すると、図 3(c) の状態は起こらずに PE の隣接関係を保持することができる。

以下に動的負荷分散法の手順を示す。

1. 各 PE は、8 近傍 PE および自分自身を含む 9PE 中で、前のタイムステップの計算時間が最小の PE を求め、送り先 PE とする。
2. 送り先 PE が図 4 に示した send PE または receive PE にあたるときは、次のいずれかの手順を行なう。
  - 送り先 PE が send PE にあたるとき、分子の数が非零かつ最小のセルを送る

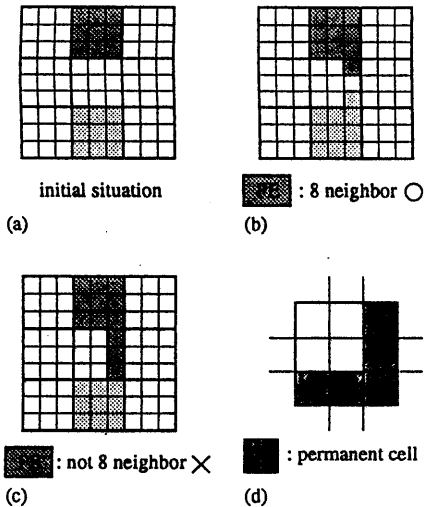


図 3: セルの再分配の概要 (a) 初期状態 (b) セルの再分配 (成功例) (c) セルの再分配 (失敗例) (d) 固定分配セル

- 送り先 PE が receive PE にあたるとき、既にその PE から受けとったセルがあれば、その中で分子の数が非零かつ最小のセルを送る

送り先 PE が send PE、receive PE のどちらにも該当しない場合はセルの転送は行なわない。周囲 PE からセルを送られるとき、それを受けとる。

この一連の操作は、分子の再分配(運動して他のセルに移動した分子のデータを正しいセルに帰属させる)の前に行なう。このとき各 PE は、移動するセルの番号および移動先 PE の番号だけを全 PE にブロードキャストして、セルの移動を行なう。実際に分子のデータを移動するのは、分子の再分配を行なうときである。

今回は、分子データの移動に同期通信を用いた。一般に同期通信では、データ転送時間に対して通信開始時間が非常に大きい。今回行なったシミュレーションの範囲では、データ転送時間よりも、通信開始時間のオーバーヘッドが大きかった。そのため、セルの再分配で分子データを転送する通信、および分子の再分配で分子データを転送する通信をまとめることによって、通信時間を小さくすることができる。

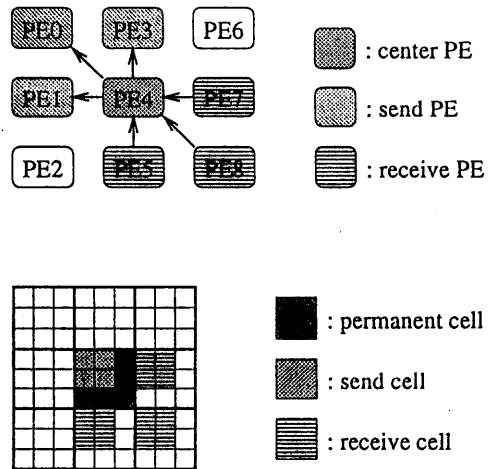


図 4: 動的負荷分散法の概要

#### 4 動的負荷分散法の性能評価

CGC 法を、64PE を有する並列計算機 CM-5 上 に実装した。動的負荷分散を行なわない CGC 法の、1PE に対する高速化率を図 5 に示す。分子の数は 5832、セル数 512 に固定し、物理条件は [1] で用いられている組み合わせとした。時間計測には高温の状態を用いたため、負荷の不均衡は小さく、計算時間にほとんど影響しないと考えられる。今回使用した PE 数の範囲では、使用 PE 数程度の高速化が得られた。そのため、CGC 法は 64PE 程度を使用する際には有効であると考えられる。

次に、動的負荷分散を行なうために処理時間を詳しく調べる。分子動力学法では一般に、相互作用の計算時間が全体の処理時間の大きな割合を占める。相互作用の計算時間は、PE 間の負荷の不均衡が問題になる部分でもある。負荷の不均衡が計算時間におよぼす影響を調べるために試験計算を行なう。着目する量は、1 タイムステップ中で、全体の計算時間と相互作用の計算時間である。1 タイムステップ全体の計算時間については、全 PE 中で最大値と最小値を調べた。ただし、各タイムステップの終わりに全 PE で同期をとるため、最大値と最小値の差異は小さいと予想される。相互作用の計算時間の幅を調べるために、1 タイムステップごとに、全 PE における相互作用の最大計算時間および最小計算時間を調べた。

試験計算は、CGC 法プログラムを用いて、低温、低密度の状態で分子が時間とともに集中するシミュレーションで行なった。2744 分子、512 セルのシミュ

レーションの各タイムステップでの計算時間変化を  
図 6 に示す。グラフで示した量は以下の通りである。

**time.max** 1 タイムステップの最大計算時間

**time.min** 1 タイムステップの最小計算時間

**force.max** 相互作用の最大計算時間

**force.ave** 相互作用の平均計算時間

**force.min** 相互作用の最小計算時間

図 6 では、以下のことが観察される。

- 1 タイムステップの計算時間は、最大値と最小値の 2 本の線であるが、差が非常に小さいために 1 本の線に見える
- 力の計算時間は、時間とともに最大値と最小値の幅が増大する
- 1 タイムステップの計算時間は、相互作用の最大計算時間に依存する

そのため、このようなシミュレーションでは、動的  
負荷分散法が必要となる。

次に、動的負荷分散法を CGC 法に組み合わせて  
実装を行なった。動的負荷分散法の有効性を調べる  
ために図 6 と同様の物理条件で計算時間を比較した。  
64PE を用いた 17576 分子、4096 セルのシミュレ  
ーションの、各タイムステップでの計算時間変化を図 7  
に示す。ここで、図中の no\_load\_balance は CGC 法、  
load\_balance は動的負荷分散法を用いた計算時間  
である。また、400、1000、1600 の各ステップでデー  
タの書き出しを行ない、一度プログラムを終了する。  
書き出したデータを読み込んでプログラムを再開し、  
シミュレーションを継続している。動的負荷分散法  
のプログラムでは、CGC 法で書き出したデータを読  
みこんでいるために、計算時間にとびが生じている。

シミュレーションの初期は比較的計算負荷が均等  
であるために、動的負荷分散法のオーバーヘッドが  
大きい。タイムステップが進むと分子の集中が顕著  
となり、動的負荷分散法を行なわない場合は急激に  
処理時間が増加する。しかしさらにタイムステップ  
が進むと分子が極端に集中するために、動的負荷分  
散法を使用しても、計算負荷を分散しきれなくなっ  
ている。これは、動的負荷分散法で固定分配セルを  
定義したために、セルの再分配が著しく制限された  
ことが原因と考えられる。

speed-up rate

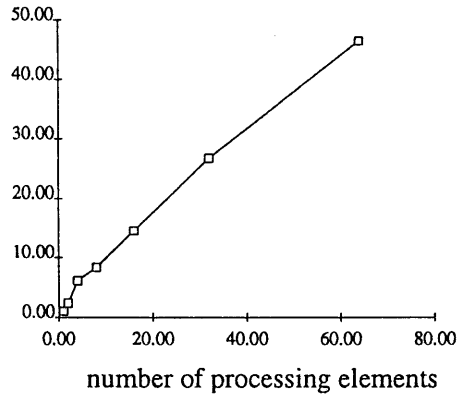


図 5: 1 プロセッサに対する CGC 法の高効率化率

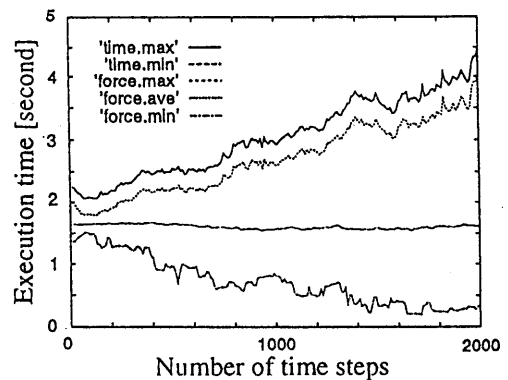


図 6: CGC 法の計算時間変化

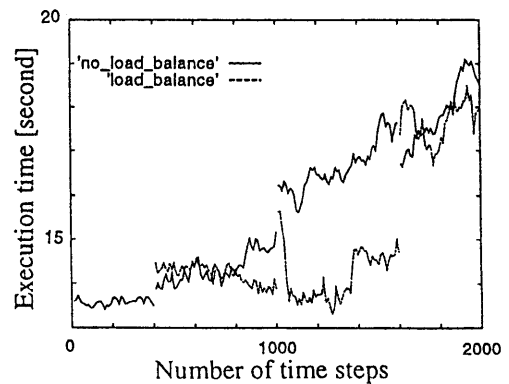


図 7: CGC 法と動的負荷分散法の計算時間変化

## 5 まとめ

本論文では、CGC法を用いた分子動力学法の並列化アルゴリズムの動的負荷分散法を提案し、その有効性について検討した。提案した動的負荷分散法は、セルを単位としたタスクスケジューリングを用いた比較的簡単な方法であるが、CGC法において計算負荷を平均化できた。しかし、希薄な気体が完全に液体になる時など、極端に分子が集中する時には十分な負荷分散が行えなかった。

今後は負荷分散を行なうしきい値を導入し、負荷分散効果を改善する手法について議論する予定である。

## 参考文献

- [1] D. W. Heermann, (小澤哲、篠嶋妥 訳)、シミュレーション物理学、シュプリンガー・フェアラーク東京、1990
- [2] T. Hoshino, K. Takenouchi, *Processing of the molecular dynamics model by the parallel computer PAX*, Computer Physics Communications 31(1984)287-296
- [3] D. M. Beazley and P. S. Lomdahl, *Message-passing multi-cell molecular dynamics on the Connection Machine 5*, Parallel Computing 20(1994)173-195.
- [4] F. Hedman and A. Laaksonen, *An approach to data parallel molecular dynamics for liquids*, International Journal of Modern Physics C Vol.4, No.1 (1993) 41-48
- [5] P. Tamayo, J. P. Mesirov and B. M. Boghosian, *Parallel Approaches to Short Range Molecular Dynamics Simulations*, Proceedings of Supercomputing 91(1991).
- [6] F. Bruge, S. L. Fornili, *Concurrent molecular dynamics simulation of spinodal phase transition on transputer arrays*, Computer Physics Communications 60(1990)31-38
- [7] *Connection Machine CM-5 Technical Summary*, Thinking Machines Corporation, November 1992