

カーボンクラスター生成並列分子動力学法シミュレーションの 高速化手法

林 亮子 樋渡 保秋† 堀口 進

北陸先端科学技術大学院大学 情報科学研究科

† 金沢大学 理学部 計算科学科

本稿では、カーボンクラスター生成過程を模擬する分子動力学法シミュレーションの並列化性能について報告する。この並列シミュレーションで使用する相互作用は短距離相互作用であるため、領域分割に基づいたドメイン分割法による並列化手法を使用する。本稿では、まず逐次処理によって、ドメイン分割法で使用するセルの最適な大きさを決定した。カーボンクラスター生成過程では、2原子の結合関係からクラスターを構成する原子を決定し、温度制御などのクラスター処理を行なう必要がある。クラスター処理の並列化には2つの手法が考えられたので、それらを実装し、性能評価によって比較した。さらに、200万ステップのカーボンクラスター生成過程シミュレーションを並列実行した結果を示し、実行時間の変動について議論する。

A Molecular Dynamics Simulation for Carbon Cluster Formation on a Parallel Computer

Ryoko Hayashi Yasuaki Hiwatari† Susumu Horiguchi

School of Information Science, Japan Advanced Institute of Science and Technology

† Department of Computational Science, Faculty of Science, Kanazawa University

This report addresses a parallel simulation of classical molecular dynamics for making clear the formation process of carbon clusters. Carbon atoms in the simulation interacts with each other via one of the short-range interactions: Tersoff potential, thus our parallelization is based on Domain Decomposition Method. As the first step of parallelization, the optimum size of cell is estimated by the executions of serial program. Next, cluster processing like cluster recognition and temperature control is parallelized. We implemented two parallel cluster processing procedures and their performance are compared. At last, we executed a 2 million steps of molecular dynamics simulation and discussed the fluctuation of execution time. As the result, we found two problems: the one is the computational imbalance, and the other is the fluctuation of optimum size of cell locally.

1 はじめに

分子動力学法 (MD) シミュレーションは、材料科学の分野における物質・材料設計の有力な手段として注目されている。MD は物質を構成する分子や原子一つ一つの運動を模擬することによって物質の状態の時間変化を調べるシミュレーション手法である [1]。より複雑な物質や詳細な現象を扱うために、並列処理による MD の高速化について多くの研究が行われている [2][3][4]。しかし、低密度で不均一な物理状態のシミュレーションに応用した例は少ない。本稿では、これまで調べられてきた並列化手法を、物性分野の問題に応用した場合の性能評価結果を示す。そして、低密度で不均一な物理状態を高速に模擬す

るための要件を明らかにする。

炭素原子が籠状に結合してできるフラレン分子 [5] は近年理論と応用の両方の観点から注目され、その生成過程を解明する研究が盛んに行なわれている。数個~100個の炭素原子が結合してできるカーボンクラスターは、フラレン生成の初期段階と考えられる。丸山と山口はターソフ・ポテンシャル [6] の働く 200 原子を用いた古典分子動力学法シミュレーションを実行し、フラレン分子の生成過程を調べた [7]。このポテンシャルを使用して、彼らは最終的に完全な C_{60} のフラレン形状を得ている [8]。田中はターソフ・ポテンシャルを用いて 1000 原子のシミュレーションを行なって密度と原子配置の関係を調べ、ナノ秒程度のシミュレーションでクラスターが生成する

密度を評価した [9]。しかし、フラレン程度の大きさのクラスターが生成する物理条件および生成過程は十分調べられていない。クラスターの生成過程を調べるためには原子数千個以上の大規模シミュレーションが必要であり、複数のプロセッサ要素 (PE) を持つ並列計算機が有効である [10]。本稿では、カーボンクラスター生成を扱う分子動力学法シミュレーションを並列化し、高速化した結果を示す。

本稿の構成は以下の通りである。第 2 章は、本稿で扱う MD シミュレーションおよびドメイン分割法の概要を示す。第 3 章は並列分子動力学法において考慮した諸問題についての議論である。4 章では、逐次処理の実行時間と並列実行時間を比較し、その性能を議論する。5 章はまとめである。

2 並列分子動力学法

2.1 シミュレーションの概要

本稿で扱う MD シミュレーションは古典 MD と呼ばれ、各分子に働く力を分子間相互作用から計算してニュートンの運動方程式に代入し、数値積分を行うことにより時間刻み (タイムステップ) ごとの各分子の位置と速度を計算する。分子間相互作用は多くの場合、分子間距離に依存するポテンシャル関数で定義する。本稿では、このポテンシャル関数にターソフ・ポテンシャル [6] を使用する。ターソフ・ポテンシャルは相互作用の影響範囲が有限な、短距離相互作用である。そのため、扱う原子の数を N とすると、相互作用の局所性を利用して計算複雑度を $O(N)$ にすることができる。

丸山ら [7] および田中 [9] の研究成果により、本稿では密度 0.1nm^{-3} 、温度を 3000K としてシミュレーションを行なう。1 タイムステップは 0.407 フェムト秒、つまり 0.407×10^{-15} 秒である。ターソフ・ポテンシャルの具体的な表式とパラメータの数値の詳細は付録に付す。実験的には、フラレンはヘリウム、アルゴンなどの希ガス中でそれらの分子と衝突し、運動エネルギーを失いながら成長する。MD で希ガス分子まで扱うのは実行時間の面から困難である。本稿では、クラスターごとの温度制御により、クラスターと希ガス分子の衝突を簡略化して扱う。クラスターごとの温度制御は、クラスターの回転、振動、並進の運動エネルギーから得られる温度に係数を乗算して一定にする。ここではクラスターごとに計算

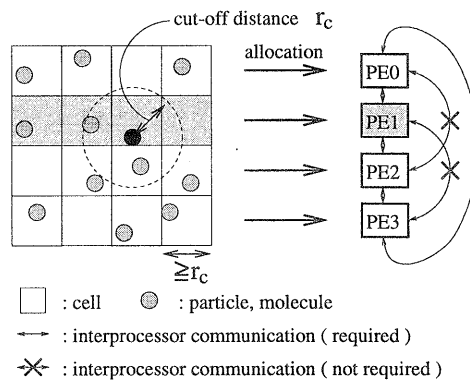


図 1: ドメイン分割法

した 3 つの温度をそれぞれ 3000K に保存した。温度制御は 100 タイムステップごとに行なった。

2.2 ドメイン分割法の概要

ドメイン分割法は、短距離相互作用の働く MD シミュレーションの並列化手法の一つである [2]。 N 個の分子を P 個の PE を用いて計算するとき、計算複雑度を $O(N/P)$ にすることができる。

ドメイン分割法の概要を示す。図 1 は、ドメイン分割法を 2 次元空間の MD シミュレーションに適用した例である。図 1 では正方形のセルを使用して、セルの大きさを相互作用影響範囲 r_c よりも大きく定める。そして 1 次元トラス結合の 4PE を使用し、1 行のセルを 1 つのドメインとして PE に割り当てる。ドメインを使用することで、通信量を削減することができる。セルの大きさが r_c よりも大きいため、分子間相互作用を計算するべき分子の組み合わせは、各セル自身と 8 近傍セル中に含まれる。そのため、図 1 では各 PE は 2 近傍の PE とだけデータを交換すればよい。3 次元空間のシミュレーションの場合はセルは立方体とし、セルの大きさを力の切断距離よりも大きく定める。このとき、各セル自身と 26 近傍セル中の分子を参照すれば、相互作用の計算に必要な分子の組み合わせが得られる。

3 並列シミュレーション

3.1 計算機環境

性能評価のために、1 ノードあたり 1PE を持つ 128 ノードの分散メモリ型並列計算機である T3E を

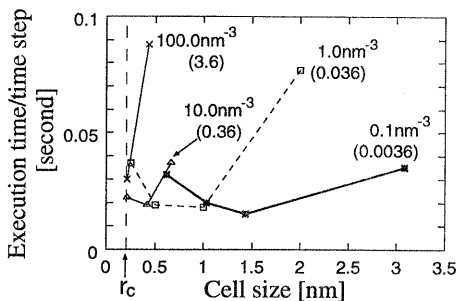


図 2: セルの大きさと実行時間の関係 (T3E の 1PE, 1000 原子, 3000K)

使用した。T3E は PE に DECchip 21164(300MHz, 600MFLOPS, 1200 MIPS) を使用している。DECchip 21164 はそれぞれ 8KB のデータキャッシュと命令キャッシュ, および 96KB の 2 次キャッシュを持つ。結合網は 3 次元トーラス型であり, 通信性能はノードあたり実効 2.8GB/second である。

並列 MD シミュレーションの実装には, プログラミング言語 FORTRAN90 にメッセージ通信ライブラリ MPI(Message-Passing Interface) を併用し, SPMD 型処理を行った。本稿で用いる実行時間は, MPI の MPLWTIME 関数を使用して計測した。比較のための逐次処理では, T3E の 1 ノードを使用し, 逐次処理プログラムを実行した。実行時間は, 並列処理と同様に, MPI の MPLWTIME 関数を使用して計測した。

3.2 並列分子動力学法における諸問題

3.2.1 最適なセルサイズ

本稿では, セルの個数だけループを繰り返すようにドメイン分割法を実装する。このとき非常に希薄な系では, セルを大きくするとループの繰り返し数が小さくなるために, 実行時間を削減可能である。一方, セルを大きくすると相互作用をしない原子の組み合わせについても原子間距離を計算するために, オーバーヘッドが生じる。

最適なセルの大きさを決定するために, 逐次処理を用いてセルの大きさと実行時間の関係を 4 種類の密度について調べた結果を図 2 に示す。これらの計算は, 初期状態付近, つまり原子が一様に分布した状態で実行した。相互作用影響範囲中に何個の原子が

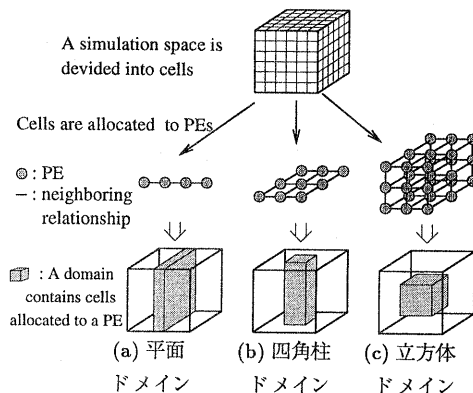


図 3: 3 次元 MD シミュレーションの 3 種類のドメイン形状

あるかは次式で表せる。

$$\frac{4}{3}\pi r_c^3 \approx 0.036\rho, (r_c = R^{(2)} = 0.2nm). \quad (1)$$

各密度での式 1 の数値を図 2 中に括弧で示した。図 2 によると, 密度が小さくなるほど, 実行時間が最小となるセルの大きさは大きくなる。図 2 によれば, 密度 $0.1nm^{-3}$ の時の最適なセルの大きさは $1.4nm$ である。

図中の式 1 の数値によると, 密度が $10.0nm^{-3}$ 以下ではいずれの密度でも, ほとんどの原子は他の原子と相互作用をしていないことがわかる。通常ドメイン分割法ではセルの大きさは 1 種類であり, ある密度で最適なセルの大きさが, 他の密度では実行時間を増加させると考えられる。以上のことから, 相互作用の計算量が増えない場合でも, 密度が不均一になるだけで計算時間が変動することが予測できる。

3.2.2 最適なドメイン形状

3 次元空間のシミュレーションでは図 3 に示すように, 3 通りのドメイン形状が可能である。林らはドメイン形状と通信時間の関係について考察し, 最適なドメイン形状の決定法を示した [4]。詳細は文献 [4] に譲るが, 要点は使用する計算機環境での通信時間の評価式に基づいた, 各形状での通信時間の予測である。T3E 上での最適なドメイン形状を予測した結果を図 4 に示す。今回は最大で 128PE を使用し, 10^5 個~ 10^6 個の原子を扱いたいので, 平面ドメインか四角柱ドメインが最適である。本稿では最大で 8000 個の原子を扱うので, 平面ドメインを用いる。

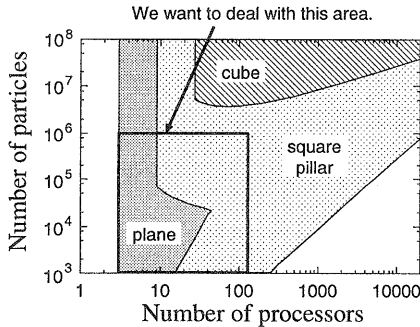


図 4: 扱う原子数および使用プロセッサ要素数と最適なドメイン形状の関係 (T3E, セル中の平均原子数が 0.27 個の場合 ($\rho = 0.1nm^{-3}$, セルの大きさが $1.4nm$ の場合に相当))

3.3 クラスター処理の並列実装

クラスターごとの温度制御を行なうためには、結合している 2 つの原子の対をもとにクラスターを構成する原子を辿り、クラスターごとに重心や速度などを計算する。これを本稿ではクラスター処理と呼ぶことにする。クラスター処理は深さ優先探索問題である。並列化にあたって 2 種類の実装を行なった。

逐次版クラスター処理 全ての原子の位置と速度を 1 個の PE に一旦集め、逐次処理的に処理する。全ての組み合わせについて 2 個の原子間距離を計算するので、計算複雑度は $O(N^2)$ である。

並列版クラスター処理 相互作用の計算部分で 2 原子の結合関係を調べておき、深さ優先探索を並列化してクラスター処理を行なう。各 PE 内でクラスターの部分的な計算を行ない、PE 間通信を行なって部分クラスターの結合関係を調べてクラスター全体の物理量を計算する。計算複雑度は $O(N/P)$ である。

逐次版クラスター処理は計算複雑度の面から不利であるが、実装が容易であるため、並列化の初期段階で実装した。

4 並列処理性能

4.1 逐次処理との比較

同じ物理条件で逐次処理、逐次版クラスター処理による並列処理、並列版クラスター処理による並列

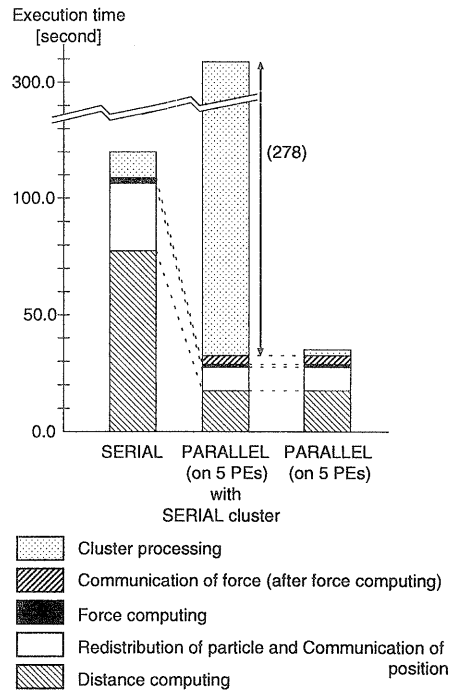


図 5: 逐次処理および並列処理によるシミュレーション実行時間とその内訳 (4913 原子, 2000 タイムステップ, $\rho = 0.1nm^{-3}$, 3000K)

処理を用いた MD シミュレーションの T3E 上での実行時間を図 5 に示す。図 5 によると、並列処理により MD の中核部分、すなわち原子間距離の計算および原子の再分配の処理で処理時間を削減可能であることがわかる。ただし、逐次版クラスター処理は実行時間を非常に大きく増大させる。この処理時間のほとんどが原子間距離の計算であり、 $O(N^2)$ であることが実行時間を悪化させている。一方並列版クラスター処理では、逐次処理と逐次版クラスター処理よりもクラスター処理時間を削減可能であり、 $O(N/P)$ であることから PE 数を増加可能である。そこで、以後は並列版クラスター処理を採用するものとする。

図 5 によると、5PE を使用した並列処理により、逐次処理の実行時間を $1/3 \sim 1/4$ に削減可能である。並列処理では PE 間通信およびその前後の処理が必要であり、その分の実行時間が増加するため、本稿で行なった並列化は満足すべき並列化として良い。

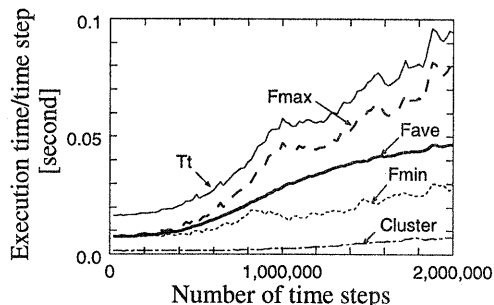


図 6: 並列処理によるシミュレーション実行時間のタイムステップによる変化 (10PE 使用, 8000 原子, $\rho = 0.1\text{nm}^{-3}$, 3000K)

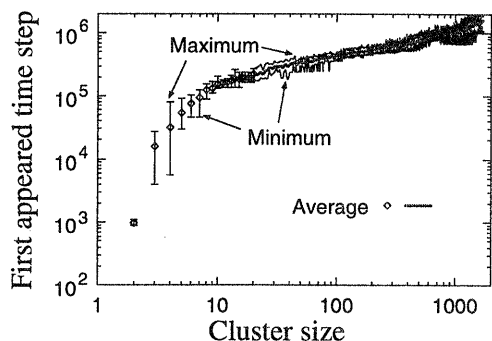


図 7: クラスターの大きさとそのクラスターが最初に出現するタイムステップとの関係 (8000 原子, $\rho = 0.1\text{nm}^{-3}$, 3000K, 10 回のシミュレーションの平均)

4.2 計算負荷の時間変動

次に、カーボンクラスター生成シミュレーションの進行と実行時間変化の関係を図 6 に示す。図 6 中の量は以下の通りである。

Tt 1 タイムステップの並列処理時間

Fmax 10PE における最大の相互作用計算時間 (原子間距離の計算を含み、Tt の内数である。Fave と Fmin も同様である。)

Fave 10PE における平均の相互作用計算時間

Fmin 10PE における最小の相互作用計算時間

Cluster クラスター処理の最大計算時間

Fmax と Fmin の差は計算負荷の不均衡の大きさを表し、Fave は理想的に計算負荷を分散できた場合の相

互作用の計算時間を表す。図 6 によると、シミュレーションの進行とともに計算負荷の不均衡が生じ、Tt を増加させていることがわかる。さらに、Fave は最終的には初期状態の 5 倍程度になり、増加が顕著である。

このシミュレーションにおけるクラスター生成状況を示すのが、図 7 である。図 7 によると、100 万ステップ付近で 1000 個以上の原子からなるクラスターが出現していることがわかる。原子が結合したと判定する距離は 1.8\AA であることから、この状況では、局所的に最大で 100nm^{-3} 程度の密度になっていることが予想される。このとき、セルの大きさは 1.4nm であり、相互作用影響範囲 0.2nm よりも非常に大きいため、原子間距離計算のオーバーヘッドが大きくなっていると予想される。Fave の増加は、相互作用計算が生じていることに加えて、セルの大きさが適切でないことの効果があると考えられる。

4.3 並列計算の課題

本稿で扱った MD シミュレーションは短距離相互作用であり、これまでに研究が進んでいる短距離相互作用の並列化手法で扱うことができる。ただし、これまでに並列化されてきたシミュレーションと異なり、非常に希薄な系であるために、並列化にあたっては並列化手法の実装の詳細を再検討する必要がある。例えば、高速な実行のためにはシミュレーション空間を適切な大きさに分割する必要がある。クラスター生成過程では、局所的に原子の密度にゆらぎが生じることから、計算負荷の平均化問題が生じる。さらに、特に希薄な系では、最適なセルサイズの局所の変動が実行時間を増加させる。並列計算において、さらなる高速化を達成するためにはこれらの課題を解決する必要がある。

5 まとめ

本稿では希薄かつ不均一な系であるカーボンクラスター生成 MD シミュレーションの並列化性能を報告した。本稿の物理状態では、計算負荷の不均衡が実行時間を増大させる。さらに、部分的な密度変化によって最適なセルの大きさが変化することが、実行時間を増大させることがわかった。今後は動的負荷分散法の実装と、局所的にセルの大きさを変化させる手法について検討する予定である。

表 1: ターソフ・ポテンシャルのパラメータ

| | | | | | |
|-----------|-------|-----|-----------|----------|---|
| $D^{(e)}$ | 6.325 | eV | $R^{(2)}$ | 2.0 | Å |
| S | 1.29 | | δ | 0.80469 | |
| β | 1.5 | 1/Å | a_0 | 0.011304 | |
| $R^{(e)}$ | 1.315 | Å | c_0 | 19 | |
| $R^{(1)}$ | 1.7 | Å | d_0 | 2.5 | |

謝辞

金沢大学 田中健児氏 (現在 (株) エプソンソフト開発センター) には、MD シミュレーションプログラム作成にあたって尽力を戴いた。また、本研究の一部は、北陸先端科学技術大学院大学 教育改善推進費 助手研究促進経費を用いて行なわれた。関係各位に感謝する。

付録：ターソフ・ポテンシャル

ターソフ・ポテンシャルの具体的な表式を以下に示す。

$$E_b = \frac{1}{2} \sum_i^N \sum_{j(\neq i)}^N [V_R(r_{ij}) - B_{ij}^* V_A(r_{ij})],$$

$$V_R(r_{ij}) = f_{ij}(r_{ij}) \frac{D^{(e)}}{S-1} \exp(-\sqrt{2S}\beta(r_{ij} - R^{(e)}),$$

$$V_A(r_{ij}) = f_{ij}(r_{ij}) \frac{D^{(e)}S}{S-1} \exp(-\sqrt{2/S}\beta(r_{ij} - R^{(e)}),$$

$$f_{ij}(r_{ij}) = \begin{cases} 1, & \text{for } r_{ij} < R^{(1)} \\ \frac{1}{2} \left(1 + \cos \frac{\pi(r_{ij} - R^{(1)})}{R^{(2)} - R^{(1)}} \right), & \text{for } R^{(1)} < r_{ij} < R^{(2)} \\ 0, & \text{for } R^{(2)} < r_{ij} \end{cases}$$

$$B_{ij}^* = \frac{B_{ij} + B_{ji}}{2},$$

$$B_{ij} = \left(1 + \sum_{k(\neq i,j)} [G_c(\theta_{ijk}) f(r_{ik})] \right)^{-6},$$

$$G_c(\theta) = a_0 \left(1 + \frac{c_0^2}{d_0^2} + \frac{c_0^2}{d_0^2 + (1 + \cos\theta)^2} \right).$$

これらの式は多くのパラメータを含んでいる。本稿で使用したパラメータの数値を表 1 に示す。これは、カーボンクラスターの MD を実行して完全な C_{60} の形状を得た文献 [8] で使用している数値である。

参考文献

[1] Heermann, D. W.: “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] D. M. Beazley, P. S. Lomdahl, N. Grønbech-Jensen, R. Giles, and P. Tamayo: *Parallel Algorithms for Short-range Molecular Dynamics*, Annual Reviews in Computational Physics, Vol. 3, World Scientific, pp. 119–175 (1995).
- [3] 松原正純, 板倉憲一, 朴泰佑: 超並列計算機 CP-PACS における大規模分子動力学法シミュレーション, 情報処理学会論文誌, Vol. 40, No. 5, pp. 2172–2182 (1999).
- [4] R. Hayashi and S. Horiguchi: Domain Decomposition Scheme for Parallel Molecular Dynamics Simulation, *Proceedings of HPC Asia '97*, pp. 595–600 (1997).
- [5] H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl and R. E. Smalley: C_{60} : Buckminsterfullerene, *Nature*, Vol. 314, No. 14, pp. 162–163 (1985).
- [6] J. Tersoff: Modeling solid-state chemistry: Interatomic potentials for multicomponent systems, *Phys. Rev. B*, Vol. 39, No. 8, pp. 5566–5568 (1989).
- [7] 山口康隆, 丸山茂夫: フラーレン生成過程の分子動力学 (第 1 報, ケージ構造の形成と制御温度), 日本機化学会論文集 (B 編), Vol. 63, No. 611, pp. 2398–2404 (1997).
- [8] 丸山茂夫, 山口康隆: フラーレン生成過程の分子動力学 (第 2 報, 完全な C_{60} へのアニーリング), 日本機化学会論文集 (B 編), Vol. 63, No. 611, pp. 2405–2412 (1997).
- [9] 田中 健児: カーボンクラスターの気相からの成長の分子動力学 (1999). 金沢大学 修士論文.
- [10] Ryoko Hayashi, Kenji Tanaka, Katsuya Nakano, Susumu Horiguchi, and Yasuaki Hiwatari: A Molecular Dynamics Simulation of the Carbon Cluster Formation on a Parallel Computer, *Proceedings of the 5th International Conference on Computational Physics (Progress of Theoretical Physics Supplement No. 138, 2000)*, pp. 262–263 (2000).