

定数ラウンドの MapReduce の計算能力に関する研究

今澤 規真^{††} 和田 幸一[†]

† 法政大学理工学部

†† 法政大学大学院 理工学研究科

1. はじめに

近年, MapReduce の理論モデルである MRC[1]によって,多くの複雑性クラスの問題が解かれている. 本稿では, 効率の良い並列計算を行うことができる問題のクラスとして認知されているクラス NC と MRC モデルとの関係性について研究を行った.

2. MapReduce モデル(MRC)

本研究で使用した MapReduce 計算モデルについて定義する[1].

定義 1

mapper μ は RAM であり, key 値 $\langle k, v \rangle$ を入力として, key 値対, $\langle k_1, v_1 \rangle, \langle k_2, v_2 \rangle, \dots$ を出力する.

reducer ρ は RAM であり, 各 key k とその値のリスト $\langle v_1, \dots, v_n \rangle$ を入力として受け取り, 各 key k と新しい値のリスト $\langle v'_1, \dots, v'_n \rangle$ を出力する.

MRC マシンの実行動作は以下の通りである.

1. U_{r-1} は最後のラウンドからの key 値対のリストを示す. mapper μ_r は U_{r-1} の key 値対を入力とする. ($r=1$ の場合は, 入力の key 値対を指す.)

2. shuffle and sort は key によって値をグループ化する.

$V_{k,r} = \{k, (v_{k,1}, v_{k,2}, \dots)\}$ とする.

3. reducer ρ_r は各 $V_{k,r}$ を割り振り, $U_r = U_k \rho_r(V_{k,r})$ となる.

上記の各 r については, map ステップ, shuffle ステップ, reduce ステップの 3 つから構成される.

MRC とは MapReduce Class の略称であり, 以下の条件を満たす.

MRCⁱ は mapper μ_r と reducer ρ_r で構成される列 $\langle \mu_1, \rho_1, \mu_2, \rho_2, \dots, \mu_R, \rho_R \rangle$ であり, μ_1 の出力が ρ_1 の入力となり, ρ_1 の出力が μ_2 の入力となる. また ρ_R の出力は MRCⁱ の最終的な出力となる. RAM である mapper と reducer は容量が $O(N^c)$ であり, それぞれ $O(N^c)$ 個ずつ存在する. N は入力サイズで, c は $\frac{1}{2} \leq c < 1$ を満たす. (本論文では $c = \frac{1}{2}$ に固定している) また, $R = O(\log^i N)$ である.

3. 結果

複雑性クラス NC^1 (深さ $O(\log^k n)$, 多項式サイズ, ファンインが定数に制限された AND, OR, NOT ゲートからなる論理回路によって解ける問題のクラス) に属する定数幅の置換分岐プログラム w-PBP を, 9 ラウンドの MRC でシミュレートすることに成功した.

アルゴリズムの大まかな概略を図 1 に示す.

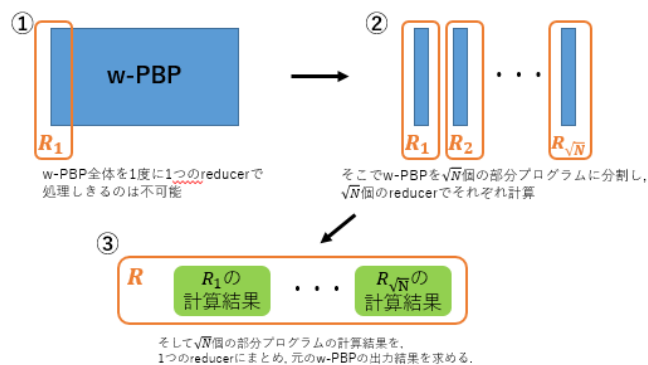


図 1 w-PBP を MRC でシミュレートする流れ

reducer の容量では, 置換分岐プログラム全てを計算することはできないため, 置換分岐プログラムを \sqrt{N} 個に分割し, 分割した部分プログラムを各 reducer 内で計算し, その結果を 1 つの reducer に集め, 計算することでシミュレートを行った.

4. 今後の課題

現状のラウンド数をさらに減らすことができないか, また NC^k の場合について MRC の関係性の解明などが今後の課題である.

参考文献

[1] Karloff, H., Suri, S., Vassilvitskii, S. "A model of computation for mapreduce" SODA 2010, pp. 938-948. Society for Industrial and Applied Mathematics, Philadelphia (2010)