

Секция «Биоинженерия и биоинформатика»

Поиск потенциальных возможностей распараллеливания алгоритма множественного выравнивания нуклеотидных и белковых последовательностей ClustalW2

Пироженко Илья Сергеевич

Студент

*Филиал МГУ имени М.В.Ломоносова в г. Севастополе, Факультет компьютерной
математики, Севастополь, Украина*

E-mail: ilya.pir@gmail.com

Одним из краеугольных камней биоинформатики является сравнение или выравнивание нуклеотидных и белковых последовательностей. Множественное выравнивание биологи используют для поиска диагностических моделей, установления семейств белков, нахождения гомологии между новыми и уже существующими последовательностями, как необходимую часть более сложных анализов. Таким образом, задача создания эффективного и универсального алгоритма множественного выравнивания является важной составляющей развития биологии.

Существует довольно много последовательных алгоритмов множественного выравнивания, таких как: T-Coffee, MUSCLE, MAFFT, ClustalW2. В статье используется ClustalW2 ввиду того что он: наиболее широко используется биологами по всему миру; надежен; обладает простым и понятным C++ кодом, в отличие от конкурентов; работает на всех основных платформах; выдает точные результаты [1].

Однако алгоритм ClustalW2 не имеет параллельной реализации. Активное внедрение многопроцессорных и многоядерных архитектур дает возможность значительно сократить время счета программы, тем самым ускорив решение биологических задач. В статье производится поиск возможностей по распараллеливанию алгоритма ClustalW2, дается теоретическая оценка выигрыша производительности.

Литература

1. Kuo-Bin Li. ClustalW-MPI: ClustalW analysis using distributed and parallel computing. //Bioinformatics Vol. 19, No. 12, 2003, pp. 1585-1586.
2. Yongchao Liu, Bertil Schmidt, Douglas L. Maskell. Parallel Reconstruction of Neighbor-Joining Trees for Large Multiple Sequence Alignments using CUDA. //IEEE Computer Society, pp. 1- 8.
3. M.A. Larkin, G. Blackshields, N.P. Brown. Clustal W and Clustal X version 2.0. //Bioinformatics Vol. 23, No. 21, 2007, pp. 2947-2948.