

Solving multiple sequence alignment problem utilizing integer linear programming

Pudiahwai Anton Wibowo, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=125557&lokasi=lokal>

Abstrak

Salah satu dari masalah-masalah dominan pada komputasi biologi molekuler adalah penyejajaran barisan berganda (Multiple Sequence Alignment - MSA) dari DNA. Banyak metode yang telah diajukan untuk menyelesaikan masalah MSA seperti pemrograman dinamik dan heuristik. Satu metode telah diajukan oleh Althaus et al. untuk menyelesaikan masalah MSA yang didasarkan pada pemrograman linear bilangan bulat (Integer Linear Programming - ILP). Formulasi ILP umum dari masalah MSA diturunkan dari representasi graf dari masalah MSA. Walaupun formulasi ILP umum dari masalah MSA diketahui, membentuk model ILP dari suatu masalah MSA yang dapat diselesaikan langsung menggunakan suatu solver ILP tidaklah mudah. Sebuah program yang dapat membangun dan menyelesaikan model ILP dari sebuah masalah MSA menggunakan MATLAB telah dibuat. Metode yang digunakan untuk menyelesaikan model ILP tersebut adalah branch-and-bound. Program yang telah dibuat dapat menghasilkan model ILP dari sembarang masalah MSA yang diberikan tetapi hanya dapat menyelesaikan masalah MSA dari sejumlah kecil barisanbarisan DNA yang pendek. Hasil dari program tersebut adalah penejajaran barisan-barisan DNA dari masalah MSA yang diberikan.

<hr>

One of the dominant problems in computational molecular biology is multiple sequence alignment (MSA) of DNA. Many methods have been proposed to solve MSA problem such as dynamic programming and heuristic. A method has been proposed by Althaus et al. to solve MSA problem which is based on integer linear programming (ILP). The general ILP formulation of the MSA is derived from the graph representation of the MSA problem. Although we have the general ILP formulation of the MSA problem, constructing the ILP model of an MSA that can be solved directly using an ILP solver is not straightforward. We develop a MATLAB program that can generate and solve the ILP model of an MSA problem. The method that is used to solve the ILP model is branch-and-bound. The constructed program can generate the ILP model of any given MSA problem but can only solve an MSA problem of a small number of short DNA sequences. The result of the program is the aligned sequences of the MSA problem.