

加速演算装置を用いた相同性検索の評価

岡山理科大学 情報理工学部学部 木戸 善之

目的：生物情報を解析をする上で、遺伝子間の相同性、発現量相関を解析することは重要であると言える。本研究では、遺伝子間の相関を解析するアプリケーションSiGN-BNをGPGPU/ベクトルプロセッサで最適化することを目標としている。

内容：すでにCPU版があるSiGN-BNは、遺伝子の相同性を検索し、時系列発現データをB-splineノンパラメトリック回帰にて相関を予測する。MPIで並列処理が可能だが、GPGPUやベクトルプロセッサといった行列計算を行うプロセッサでの最適化ができておらず、チューニングとして時系列セグメントごとのノンパラメトリック回帰による相関係数を算出するプロセスをGPGPUで行うようプロファイリングおよびチューニングを行った。

結果：遺伝子サンプル400、時間ステップ100という小規模データで行った場合、CPUノードとGPGPUではCPUノードで行った方が速かった。サンプル数を増やすと遺伝子間相関を行う行列計算が大きくなるためGPGPU/ベクトルプロセッサを行う方がより高速に処理できると予測できる。

利用した計算機 SQUID GPCPUノード/汎用CPUノード

ノード時間	258.37 時間
使用メモリ	24 GB
並列化	1ノード ノード内 24並列

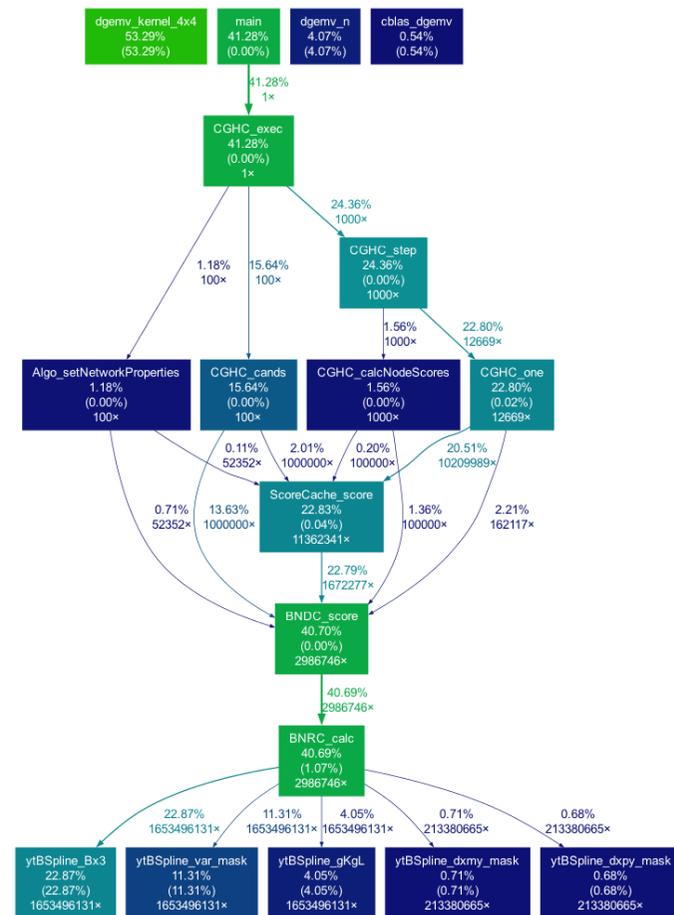


図. プロファイリングの可視化