



## 大規模分子配列データベースからの知識発見

～ 配列データマイニングを用いた配列パターン抽出とその並列分散処理～



参加費:無料(予定)

連絡先 E-mail:

ieee-smc-hiroshima-exec@smc-hiroshima.info.hiroshima-cu.ac.jp

その他:

詳細は下記の URL をご覧下さい:

<http://www.smc-hiroshima.info.hiroshima-cu.ac.jp/>

事前にコミュニティサイト

<http://goo.gl/LkHST>

で該当イベントに参加登録が必要です。

Speaker: 広島市立大学大学院情報科学研究科  
田村 慶一 准教授

Date: 2012年8月24日(金)

Time: 13:00 to 14:30

Location: 広島市立大学 情報処理センター5F

**Summary:** 分子配列データベースとは DNA の塩基配列やタンパク質のアミノ酸配列などの遺伝子配列を格納したデータベースであり、バイオインフォマティクス分野において分子配列データベース中に格納された配列データが持つ意味を解明することが重要な研究課題となっている。配列データマイニングは、配列データベース中に現れる頻出する配列パターンや類似する配列パターンを抽出する技術であり、分子配列データベースに格納されている配列データが持つ意味を解明するための基盤技術として期待されている。例えば、アミノ酸配列中に現れる特徴的な配列パターンはモチーフと呼ばれ、生物進化の過程で保存されてきた重要な生体機能に関連している。アミノ酸配列中で頻出する配列パターンはこのモチーフを含んでいる可能性が高く、配列データマイニングを用いて頻出する配列パターンを効率的に取り出すことが出来れば、モチーフ発見に寄与することができる。また、分子配列データベースに格納される配列データは指数関数的に増加しており、効率的な手法を用いたとしても非常に多くの計算時間を必要とするため、並列分散処理を用いて高速化する必要がある。本発表では、配列データマイニングとその並列分散処理について概観し、講演者が取り組んでいる分子配列データベースから頻出配列パターンと類似パターンを抽出する手法について述べるとともに、実際の大規模 PC クラスタやマルチコアプロセッサを用いた並列分散処理について述べる。



IEEE

IEEE SMC Hiroshima