

IT21 の会（平成 21 年 12 月）第 138 回議事録

日 時：平成 21 年 12 月 4 日(金) 18 時 30 分～20 時 50 分

場 所：日本技術士会 葦手第二ビル 5 階 A・B 会議室

出席者：13 名

配布資料

- ・ 0912-1 大規模データ解析の臨床への応用
The CAMDA 2009 meeting in Chicago レポート (石井氏)
- ・ 0912-2 2009 年度合宿運営プロジェクト憲章 (工藤氏)
- ・ 0912-3 合宿場所について (佐野氏)
- ・ 0912-4 CPD 中央講座プログラム (黒澤氏)
- ・ 他 CPD 行事参加票

議 事

1. 議事および資料確認 (石井氏)

2. 役員からの報告 (山下氏)

- ・ 年会費の受領通知を送付しないことにする。個人的に受領通知を望まれる方は、会計担当個人宛にその旨メールを送っていただければ、通帳を確認の上、対応する。

3. メインテーマ (石井氏)

「大規模データ解析の臨床への応用

The CAMDA 2009 meeting in Chicago レポート」

(1) ゲノム解析の最新動向

- ・ ゲノム解析にはゲノム DNA 配列解析、ゲノム機能解析 (ゲノム発現解析、非コード RNA 解析、ゲノムメチル化解析)、プロテオーム解析、メタボローム解析などがある。
- ・ 特に、ゲノム機能の解析として、メッセンジャー RNA の発現調節を調べるトランスクリプトーム (遺伝子発現解析)、マイクロ RNA 解析、メチル化解析

(エピゲノミクス)、ChIP 解析 (ゲノムに結合するタンパク質の解析) が行われる。

- これらのゲノム機能の解析に威力を発揮する機械が「マイクロアレイ」と「次世代シーケンサー」である。そこで、まずマイクロアレイと次世代シーケンサーのしくみと応用について解説した。マイクロアレイと次世代シーケンサーからは、大量のカテゴリカルデータ、テキストデータ、数値データが算出され、情報技術の利用が不可欠である。
- これらは、癌や感染症などの診断や予後の判定、新薬開発の指標、幹細胞、iPS細胞の作成とその評価、品種改良や、環境微生物の評価、食品成分のゲノムレベルでの評価(ニュートリゲノミクス)などに利用される。

(2) The CAMDA 2009 meeting in Chicago 報告

- 今回紹介した CAMDA2009 とは、Clinical Assessment of Massive Data Analysis の略で、Clinical Assessment of Microarray Data Analysis と以前は呼ばれていた。
- マイクロアレイデータのデータ解析のためのデータマイニング技術の研究会として設立され、世界中の研究者が同じデータセットを用いてのデータ解析の共同研究、コンテスト、解析技術の研究発表会などを用いて、臨床でのデータ解析、データマイニングの向上を目的とした国際会議である。
2000年に Duke 大学の生物情報関係の研究者らにより設立された。
- 今年の会議の目玉は、最新鋭のゲノム解析装置である次世代シーケンサーである。次世代シーケンサーからのデータセットを用いたデータ解析のコンテスト、研究発表会を中心に会議が展開された。また、統計解析環境 R を用いた次世代シーケンサーのデータ解析を行うチュートリアル講演が行われた。今回、これらの解析を行うためのソフトウェアの概要紹介、個々の講演者の講演内容の概要紹介などを行った。
- まとめとして、次世代シーケンサーの登場によるデータ解析の高速化に伴い情報技術の需要性がますます増加していること、及び、今後の解析技術の進歩により個別化医療、個別化ゲノミクスの実現化がますます加速化していくであろうことが強調された。
- 演者は 2010 年 1 月より、次世代シーケンサーを用いたゲノム機能解析により臨床研究(癌研究)を開始する予定である。

(3) 質疑応答

・随時、演者と聴衆間で質疑応答が交わされた。要旨は以下のとおりであった。

Q. RNA のほうが網羅的解析を行いやすいとの話だったが、今話題の新型ウィルスの解析も RNA で行われているのか？

A. 新型インフルエンザの場合は、RNA を用いた解析も行われるが、たんぱく質を用いた解析も行われている。

ゲノム機能解析の観点ではなく、ワクチンを開発する必要性から実施されているためである。

Q. 「マクロアレイ」方式において、データ解析を行いやすくするために、サンプル収集時に使用する試料の並び順を工夫しているのか？

A. 検出するプローブのデザインや、種類など各メーカーが独自に工夫をしている。ユーザーは各メーカーのプロトコールに従って解析を行い、ユーザー側で工夫、改良などは行っていない。

Q. 「次世代シーケンサー」方式を用いて一個の遺伝子の情報を収集するのに要する時間はどれほどであるか？

A. 正確な時間は知らないが、測定対象や測定機器によって異なるはずである。おそらく、それほど長くないはずである（一反応あたり数十分～数時間以内ではないか？）。解析の全行程は約 1 週間程度である。

Q. CAMDA では、データマイニング解析のコンテストが実施されているとの話だが、そのコンテストにおいて優劣を決める基準は何か？

A. 生物学的見地から有意義な情報をいかにして抽出したか（マイニングしたか）によって勝敗が決まる。時間効率性・資源効率性などの情報工学的見地（ソフトウェアの品質特性）に基づく評価もされるが、生物学的見地ほど重きは置かれていないであろう。

Q. 機能ゲノム解析では「アノテーション」（試験対象となった塩基配列が、元々の DNA のどこに位置していたか）を解析することが重要であるとの話だが、アノテーション解析に有効なアルゴリズムは何であるか？

A. 隠れマルコフモデルによる解析、ベイズ解析、クラスター解析などが用いられている。

4. サブテーマ（佐野氏、工藤氏）

「2月合宿について」

・2月合宿のプロジェクト憲章、合宿場所について説明があった後、予算追加申請があった。

1. 合宿会場補助費：現在 30,000 円である。会場費（100,000 円）、機器代（20,000 円）が見込まれるため 90,000 円を追加し、120,000 円とする。

2. 会場調査費：現在この科目はない。会場下見など調査費用として新たに5,000円を設ける。

・以下のような質疑応答があった。

Q. 宿泊費は参加者個人が負担するか？

A. そのとおりである。

Q. 以前の合宿では、年末までに参加表明した人には割引特典があった。今年実施するか？

A. しない。

・その後、予算追加承認について採決し、満場一致で可決された。

5. インターネットで聴講できる欧米大学の講義(石井氏)

・近年、ハーバード大学、MIT、スタンフォード大学、プリンストン大学など欧米のトップクラスの大学がインターネットを通じて講義を公開するようになり、無料で聴講できるようになっている。

各大学や関係機関のサイト、Youtube、iPhone などから、ストリーミングポッドキャストで、講義ビデオ、講義録音、講義のシラバス、スクリプト、講義ノートが入手できる。英語に自信がなくても、スクリプトが公開されているので、理解可能である。

・主なサイトは、以下のとおり

AcademicEarth、Open Course Ware Consortium、MIT Open Course、Stanford Engineering Everywhere (SEE)、YouTube EDU、The Vega Science Trust Videos。

6. CPD 中央講座の案内 (黒澤氏)

・以下の講座の紹介があった。

(1) 12月12日(土) 13:00~17:00 アクセス青山フォーラム

「マグネシウム文明論 ―石油に代わる新エネルギー資源―」

(2) 1月12日(火) 16:00~17:30 学士会館 202号室

新春記念講演会「宇宙開発の未来～日本と世界の月惑星探査動向～」

(3) 1月21日(木) 13:00~17:00 葺手第二ビル

「ハイブリッド車プリウスと電池の開発」

(4) 2月10日(水) 18:00~20:00 葺手第二ビル

「インターネット・セキュリティ」

(5) 2月13日(土) 13:00~17:00 機械振興会館

「新型インフルエンザと課題」

(6) 3月6日(土) 13:00~17:00 学士会館

「学協会と技術士会との連携」

(7) 4月23日(金) 18:30~20:30 葺手第二ビル

「グリーンIT」

7. 初参加者の自己紹介

- ・山下 茂雄氏
- ・和泉 明子氏

8. 久しぶりにいらした参加者の自己紹介

- ・北尾 由之氏

以上 (記載者: 山下真吾)