

船井情報科学振興財団
留学報告書

第7回：Cambridgeでの4年目について

2021年1月

Funai Overseas Scholarship 奨学生 平川奇跡
kh612@cam.ac.uk

1. はじめに

Cambridgeでの生活も無事4年目を迎えることが出来た。本報告書では、最近の活動について共有する。

2. 授業について

工学士課程の4年目は大きく分けて8 module およびプロジェクトに分かれる。4つのモジュールはそれぞれミカエル学期（10月～12月）とレント学期（1月～3月）の間に履修しなければならない。

まず今学期受講しているコースを簡単に紹介したいと思う。

モジュール内のアルファベットは、モジュールが属する特定の部門（情報工学部など）を示す。

- 4F13 *Probabilistic Machine Learning*

- ・ 統計的機械学習の基本概念を十分に理解していることを示す。
- ・ 基本的なML手法を実践的な問題に応用できるようになる。

-4F10 *Deep Learning and Structured Data*

- ・ このモジュールでは、生成モデルと識別モデルに使用できるディープラーニングの基本的な概念と、構造の形態を学ぶことを目的としている。さらに、音声や言語などの構造化されたデータを分類するためのモデルの使用についても議論する。

- 4M24 *Computational Statistics and Machine Learning*

- ・ 現代のすべての工学系科学で実用的に使用できる、厳密に設計された機械学習と計算統計学のすべての手法を支える、必要とされる統計的・数学的な概念を学生に紹介する。
- ・ データとモデルが合成された一連の困難な大規模な工学的問題に対する機械学習ソリューションを設計するために必要な高度な計算統計推論手法を紹介する。

- 4E1 Intellectual Property

- ・ 知的財産管理の学際的な性質を理解する。
- ・ モジュール中に紹介された関連する概念、フレームワーク、ツール、理論を理解し、適用できる。
- ・ 戦略的知財管理が競争上の優位性を生み出し・維持するために創出できる機会を理解する。

全学期、講義は事前録音を使用してオンラインで実施された。幸いなことに、私の専門分野（情報工学）の性質上、ほとんどの仕事がりモートでできるため、学業にはあまり影響がなかった。

ほとんどの時間を図書館で作業をすることが多かった僕にとっては、学期中にかかわらずしばらく休館せざるをえなかったのは残念だった。開館しているときでも、学生は事前予約制で、3時間以上図書館にいることは禁止されていた。

他の人々と議論する機会は限られていたが、Zoomなどのオンライン会議プラットフォームは大きな助けとなった。1年の残りの大半、ビデオ会議ツールが、学生たちとの主要コミュニケーション手段であろう。

先日イギリスで発表されたロックダウンの影響で、レント学期（1月～3月）は完全オンラインで実施され、図書館に対しては厳しい制限が課せられた。実験的で実用的なプロジェクトを行っている学生にとっては、言葉では言い表せないほどのインパクトがある。しかも、最近のロックダウンは、これまでのロックダウンよりも厳しくなっている。来期、試験やプロジェクトがどのように評価されるのかなど、不透明な部分も多い。しかし、英国では急速なワクチン接種が行われており、数ヶ月後には状況が改善すると強く願っている。

3. プロジェクト

プロジェクトのタイトルは「DNAストレージチャンネルのモデル構築」である。

プロジェクトの目的は、DNAストレージチャンネルの数学モデルを構築し、欧州バイオインフォマティクス研究所のデータを用いてそれを検証することである。

現在、デジタルデータは、そのほとんどが磁気や光媒体に保存されるのが主流となっている。デジタルデータが爆発寸前の時代を迎え、デジタルデータは日々生成され、指数関数的な速度で増加している。従来の媒体では、ビッグデジタルデータストレージの喫緊の要求を満たすことが出来ない現実があ

る。そこで注目されるのがデオキシリボ核酸(DNA)で、高密度、高複製効率、長期耐久性、長期安定性などの利点を備え、新規かつ潜在的なデータ記憶媒体として期待されている。新しいDNAデータストレージの場合、ファイルまたは読み取り可能な任意のデータは、バイナリに変換され、その後、アデニン (A)、シトシン (C)、グアニン (G) およびチミン (T) からなる DNA 配列にエンコードされる。データを保持する DNA 配列は合成され、次のデータ検索まで保存される。データが検索されると、固有のデータを持つ DNA 断片が増幅され、配列が決定され、分析される。その後、DNA ベースのデータ情報は、バイナリにデコードされ、最終的には読み取り可能な情報に変換される。

このプロセスの際、合成、シーケンス、保存中に、挿入、置換、削除などのエラーが発生する可能性がある。プロジェクトでは、ストレージチャンネルの数学的モデルを構築することで、こうしたエラーパターンをよりよく理解することを目的とする。

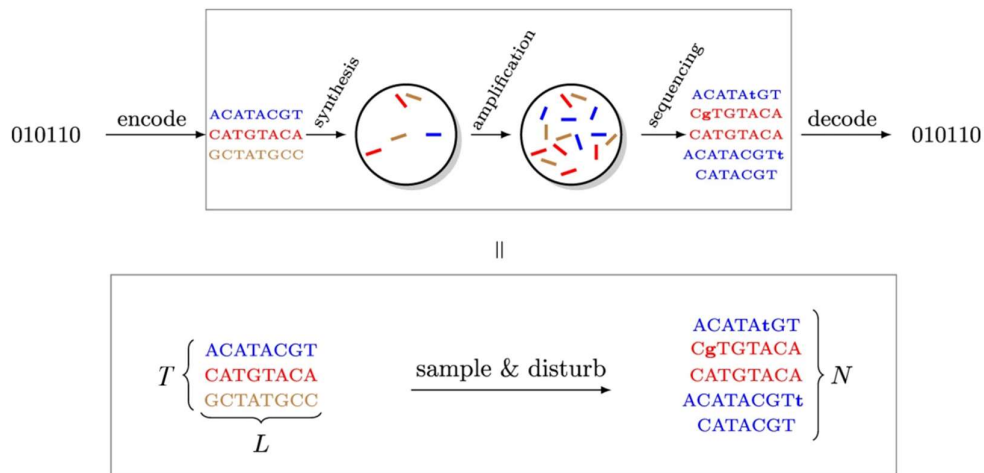


Figure 1 DNA ストレージチャンネルの構造

4. 課外活動



Figure 2 協会の方々と

ケンブリッジ大学工学協会の副会長としての役割を続けている。新型コロナウイルス感染症（COVID）関連の規制から、学会としては非常に困難な年でしたが、意外にもすぐに適応できた。

5. 最後に

ケンブリッジでの生活も残り半年となるが、常日頃から支援を頂いている船井情報科学振興財団に感謝して、この報告書の結びとさせていただく。素晴らしい環境に恵まれていることに感謝し、在学中の残された時間を充実に過ごすよう引き続き精一杯頑張りたいと思う。