

バイオ情報科学 Bioinformatics

学 年	第3学年	前期・後期	前 期	単 位	2	科 目 分 類	分子 応用 医科	選択総合
主担当教員	高須 昌子	最高評価	S	GPA	対 象			
担当教員	小島 正樹							
修得できる力	数理能力・IT力							
関連科目	1411 数学 I / 1711 生体物質学 / 1731 遺伝生化学 / 2411 統計学 / 2412 応用数学 / 2421 プログラミング基礎 / 2711 酵素学 (分子・応用) / 2711 酵素学 (医科) / 3812 進化系統学							

ねらい

バイオインフォマティクス（生物情報科学）は、生命科学と情報科学の境界領域の学問分野である。1990年代のヒトゲノム計画の進行や、その後のポストゲノム研究の流れから、大量のデータが生み出されてきた。世界的規模で蓄積された巨大なデータベースの中から、意味のある生物情報をいかに取り出し、新たな学問分野を創出するかが、今後の課題となっている。本講義では、このようなバイオインフォマティクスの概要と、その基礎となる情報科学の基本概念を取り扱う。これらの内容は、バイオインフォマティクス技術者認定試験や、経済産業省が主催するITパスポート試験、基本情報技術者試験にも役立つ。またC言語実習を行ってプログラミング能力を育成し、就職のチャンスを広げる。

一般目標

プログラミング技法と情報科学の諸概念や、バイオインフォマティクスの基礎を理解する。

到達目標

C言語の簡単なプログラムを書いて実行できる。
 2進数の簡単な計算ができる。
 論理回路の基本を説明できる。
 符号化とパリティチェックについて説明できる。
 情報量とエントロピーに関して説明できる。
 キュー、スタック、ツリー構造について説明できる。
 動的計画法を用いて、配列のアラインメントを実行できる。
 距離行列を計算して、進化系統樹を作成できる。
 アラインメントのスコアや有意性を評価できる。
 立体構造の二面角やRMSDを計算できる。
 種々のネットワークやオートマトンについて説明できる。

授業内容

回 数	担 当	テーマ	行 動 目 標
1	高須	はじめに	コンピュータシステムやいろいろなプログラミング言語があることを説明できる。情報系資格についても説明できる。
2	〃	2進数	2進数とは何か、なぜ2進数が必要かを説明できる。2進数と10進数の変換を計算することができる。C言語のプログラムのコンパイルと実行ができる。
3	〃	2進数の演算	2進数の補数と引き算の計算を行うことができる。C言語の簡単なプログラムを書くことができる。
4	〃	論理回路の基礎	論理回路の基本を説明できる。OR、AND、NOT、NAND、NOR、XORの違いを説明できる。C言語プログラミングにおいて、自分の弱点を説明できる。
5	〃	半加算回路と全加算回路	半加算回路、全加算回路を説明できる。C言語の実行エラーメッセージの内容を説明できる。
6	〃	符号化	符号化と復号化、パリティチェックについて説明できる。エラーメッセージに沿ってプログラムを修正できる。
7	〃	情報量	情報量とエントロピーに関して、説明できる。例題を見ながら練習問題のプログラムを作成できる。
8	〃	データ構造	データ構造、キュー、スタック、ツリー構造を説明できる。C言語のループを説明できる。
9	〃	ソート	ソートを説明できる。C言語の分岐を説明できる。
10	〃	前半のまとめ	前半に学んだ情報科学の基礎概念を説明できる。

回数	担当	テーマ	行動目標
11	小島	ペアワイズアラインメント	ドットマトリックス法を活用できる。配列アラインメントのスコアを計算し、動的計画法により最適化することができる。
12	〃	グラフ理論	グラフ理論の基本用語を説明できる。与えられたグラフがオイラーグラフか否かを判別できる。
13	〃	進化系統解析	配列アラインメントから距離行列を計算し、進化系統樹を作成できる。最大節約法、最尤法のアルゴリズムを説明できる。
14	〃	マルチプルアラインメント	ゲノムや配列の情報量とエントロピーを計算できる。マルチプルアラインメントの計算原理を説明できる。
15	〃	配列解析の統計学	アミノ酸の類似性を変異頻度の観点から説明できる。統計学の仮説検定によりアラインメントの有意性を評価できる。
16	〃	タンパク質の立体構造	ラマチャンドランプロットを読むことができる。フォールドクラスについて説明できる。立体構造のRMSDを計算できる。
17	〃	ネットワーク解析	ブール代数によりラクトースオペロンを解析できる。リン酸化による細胞内情報伝達機構を微分方程式系のネットワークとして解析できる。
18	〃	機械学習	与えられたデータ空間を線形分離できる。サポートベクターマシンの概要を説明できる。
19	〃	オートマトン	与えられた記号列を受理する有限オートマトンを設計できる。隠れマルコフモデル、チューリング機械、セルオートマトンの概要を説明できる。
20	〃	言語理論、小テスト	ある終端記号列が正規文法から生成されるか否かを判別できる。互いに等価な正規文法と有限オートマトンを示すことができる。パイオインフォーマティクスに関する基本的な問題を解くことができる。

準備学習：授業では実際に問題を解きながら進めるので、毎回休まず出席すれば、十分理解できる。

(予習・復習等) (高須) webに掲載されたパワーポイントで復習できる。レポートで復習できる。

(小島) 授業内容の復習とパイオインフォーマティクス技術者認定試験対策のため、期日までにCodexでCBT演習を行うこと。

授業形式：(高須) パワーポイントを用いた講義、グループワークの時間がある。演習問題を前で発表。

(小島) シナリオ学習、プリントでの作業、授業内演習(形式は同上)を併用する。授業や課題に関する連絡はCodexで行う。

課題に対するフィードバックの方法 (高須) レポートの講評を授業内で行う予定である。

(小島) CBT演習は、毎回解き終わるごとに正解と解説が表示される。小テストの正解と解説はCodexに掲載する。

成績評価方法：(高須) 授業参加度(約15%)、問題演習での貢献(約15%)、レポート(約30%)、小テスト(約40%)

(小島) 平常点(口頭試問、黒板で回答)25%、CBTのスコア33%、小テストの素点42%

教科書：(高須) 新・明解C言語入門編、柴田望洋著、2014年8月発行、ソフトバンククリエイティブ

(小島) パイオインフォーマティクス入門、日本パイオインフォーマティクス学会編、慶應義塾大学出版会

参考書：(高須) 子どもの考える力をつける 3つの秘密道具、岸良裕司、きしまゆこ著、ナツメ社、ISBN 978-4816364280

(小島) オートマトン・言語理論、富田悦次・横森貴著、森北出版

オフィスアワー：高須昌子 授業時間中または直後に質問。それ以外は担当教員とメールで打ち合わせる。

小島正樹 時間の許す限り対応(要予約) 生物情報科学教授室

教員からの一言：(高須) C言語実習は各自のペースで進めるので、2年選択科目のプログラミング基礎が未履修の方も大丈夫です。既にC言語を使える人は発展問題によって、さらにレベルアップできます。

(小島) パイオインフォーマティクスは、生命科学系の学生がデータサイエンスやAI(人工知能)の基礎を学ぶのに最適の科目だと思います。