

Universidad Católica San Pablo
Facultad de Ingeniería y Computación
Escuela Profesional de
Ciencia de la Computación
SILABO



CB309. Computación Molecular Biológica (Electivo)

2018-I

1. DATOS GENERALES

1.1 CARRERA PROFESIONAL	:	Ciencia de la Computación
1.2 ASIGNATURA	:	CB309. Computación Molecular Biológica
1.3 SEMESTRE ACADÉMICO	:	9 ^{no} Semestre.
1.4 PREREQUISITO(S)	:	CS315. Estructuras de Datos Avanzadas. (6 ^{to} Sem)
1.5 CARÁCTER	:	Electivo
1.6 HORAS	:	2 HT; 2 HP; 2 HL;
1.7 CRÉDITOS	:	4

2. DOCENTE

Mg. Jenny Linet Copara Zea

- Mag. Ciencia de la Computación, Universidad Católica San Pablo, Perú, 2017.

3. FUNDAMENTACIÓN DEL CURSO

El uso de métodos computacionales en las ciencias biológicas se ha convertido en una de las herramientas claves para el campo de la biología molecular, y éstas actualmente son usadas como parte crítica en sus investigaciones. Existen diversas aplicaciones en biología molecular relativas tanto al ADN como al análisis de proteínas. La construcción del genoma humano, por ejemplo, depende fundamentalmente de la biología molecular computacional. Muchos de los problemas de ésta área son realmente complejos y con conjuntos enormes de datos. Este curso además puede servir para ejemplificar algunos tópicos de Fundamentos de Programación (PF) y Algoritmos y Complejidad (AL) de acuerdo al Computing Curricula 2001.

4. SUMILLA

1. Conceptos Introductorios 2. Alineamiento de Secuencias 3. Clustering 4. Árboles Filogenéticos 5. Mapeo de Secuencias 6. Introducción a la Estructura de las Proteínas

5. OBJETIVO GENERAL

- Interpretar problemas biológicos haciendo uso de técnicas computacionales.
- Analizar e implementar algoritmos y estructuras aplicables al campo de la biología.

6. CONTRIBUCIÓN A LA FORMACIÓN PROFESIONAL Y FORMACIÓN GENERAL

Esta disciplina contribuye al logro de los siguientes resultados de la carrera:

- a) Aplicar conocimientos de computación y de matemáticas apropiadas para la disciplina. [Nivel Bloom: 3]
- g) Analizar el impacto local y global de la computación sobre los individuos, organizaciones y sociedad. [Nivel Bloom: 3]
- h) Incorporarse a un proceso de aprendizaje profesional continuo. [Nivel Bloom: 3]
- i) Utilizar técnicas y herramientas actuales necesarias para la práctica de la computación. [Nivel Bloom: 3]
- j) Aplicar la base matemática, principios de algoritmos y la teoría de la Ciencia de la Computación en el modelamiento y diseño de sistemas computacionales de tal manera que demuestre comprensión de los puntos de equilibrio involucrados en la opción escogida. [Nivel Bloom: 4]

7. CONTENIDOS

UNIDAD 1: Conceptos Introdutorios (0 horas)

Nivel Bloom: 3

OBJETIVO GENERAL

CONTENIDO

- Identificación de los conceptos básicos en Biología Molecular
- Reconocimiento de problemas clásicos en Biología Molecular y su representación en el campo computacional
- Aprendizaje de las herramientas de software e Internet clásicas para el campo de Bioinformática
- Introducción a los conceptos necesarios en manejo de Cadenas, Grafos y su representación algorítmica a fin de transformar problemas biológicos al tipo computacional

- Introducción a la Historia de la Genética
- Conceptos Básicos de Biología Molecular
- Problemas clásicos en Bioinformática
- Herramientas de recolección y almacenamiento de secuencias en laboratorio
- Recursos de Software, introducción a BLAST, CLUSTALW
- Cadenas, Grafos y Algoritmos

Lecturas: [Mount, 2001], [Jones and Pevzner, 2004], [Clote and Backofen, 2000]

UNIDAD 2: Alineamiento de Secuencias (0 horas)

Nivel Bloom: 4

OBJETIVO GENERAL

CONTENIDO

- Reconocimiento de las técnicas básicas usadas en el alineamiento de secuencias
- Implementación de los diversos algoritmos de comparación de secuencias
- Introducción a la programación dinámica
- Introducción y comparativa entre métodos heurísticos y exactos
- Métodos probabilísticos: PAM

- Introducción al alineamiento de secuencias
- Comparación de pares de secuencias
- Alineamiento de Secuencias Global
- Alineamiento de Secuencias Múltiples
- Cadenas ocultas de Markov
- Métodos exactos, aproximados y heurísticos del alineamiento de secuencias
- Problemas derivados del alineamiento de secuencias

Lecturas: [Mount, 2001], [Jones and Pevzner, 2004], [Clote and Backofen, 2000]

UNIDAD 3: Clustering (0 horas)	
Nivel Bloom: 4	
OBJETIVO GENERAL	CONTENIDO
<ul style="list-style-type: none"> ▪ Identificar métodos de distancia aplicables a grafos del tipo árboles ▪ Conocer la transformación de Matrices en estructuras de grafos ▪ Reconocer a los métodos de Clustering como útiles para la identificación de funciones en genes no conocidos a partir de genes similares ▪ Identificar la importancia del Clustering en el reconocimiento de patrones de enfermedades 	<ul style="list-style-type: none"> ▪ El problema del Clustering ▪ Clustering Jerárquico ▪ Algoritmo de Neighbour Joining ▪ Algoritmo del Average linkage ▪ Clustering no jerárquico o K-means ▪ EST clustering
Lecturas: [Mount, 2001], [Jones and Pevzner, 2004], [Clote and Backofen, 2000]	

UNIDAD 4: Árboles Filogenéticos (0 horas)	
Nivel Bloom: 3	
OBJETIVO GENERAL	CONTENIDO
<ul style="list-style-type: none"> ▪ Reconocer algoritmos de mediciones de distancias ▪ Analizar la complejidad computacional de cada uno de los algoritmos estudiados ▪ Reconocer la importancia de la filogenia en casos de evolución de epidemias como el HIV ▪ Utilización de herramientas de software de libre uso ▪ Implementación de los algoritmos estudiados 	<ul style="list-style-type: none"> ▪ Introducción a la Filogenia ▪ Algoritmos comunes ▪ Aplicaciones biológicas ▪ Algoritmos Exactos ▪ Algoritmos Probabilísticos
Lecturas: [Mount, 2001], [Jones and Pevzner, 2004], [Clote and Backofen, 2000]	

UNIDAD 5: Mapeo de Secuencias (0 horas)	
Nivel Bloom: 3	
OBJETIVO GENERAL	CONTENIDO
<ul style="list-style-type: none"> ▪ Identificación de problemas NP-Complejos ▪ Aplicación e implementación de técnicas diversas a fin de dar solución a éstos problemas biológicos ▪ Introducción a los métodos de tipo goloso ▪ Reconocimiento de tópicos avanzados en teoría de grafos 	<ul style="list-style-type: none"> ▪ Problema del <i>Double Digest</i> y <i>Partial Digest</i> ▪ Técnicas utilizadas en el mapeo de secuencias ▪ Mapeo con <i>Non-Unique Probes</i> ▪ Mapeo con <i>Unique Probes</i> ▪ Grafos de Intervalos ▪ Mapeo con Señales de Frecuencias de Restricción
Lecturas: [Mount, 2001], [Jones and Pevzner, 2004], [Clote and Backofen, 2000]	

UNIDAD 6: Introducción a la Estructura de las Proteínas (0 horas)	
Nivel Bloom: 2	
OBJETIVO GENERAL	CONTENIDO
<ul style="list-style-type: none"> ▪ Examina algunos tópicos de reconocimiento visual en Computación Gráfica ▪ Implementación de algunos estructuras simples como el folding 2D 	<ul style="list-style-type: none"> ▪ Fundamentos biológicos de las proteínas ▪ Motivación para la predicción de las estructuras de las proteínas ▪ Alineamiento rígido de Proteínas ▪ Técnica del alineamiento por Hashing Geométrico ▪ Predicción de la estructuras de las proteínas
Lecturas: [Mount, 2001], [Jones and Pevzner, 2004], [Clote and Backofen, 2000]	

8. METODOLOGÍA

El profesor del curso presentará clases teóricas de los temas señalados en el programa propiciando la intervención de los alumnos.

El profesor del curso presentará demostraciones para fundamentar clases teóricas.

El profesor y los alumnos realizarán prácticas

Los alumnos deberán asistir a clase habiendo leído lo que el profesor va a presentar. De esta manera se facilitará la comprensión y los estudiantes estarán en mejores condiciones de hacer consultas en clase.

9. EVALUACIONES

Evaluación Permanente 1 : 20 %

Examen Parcial : 30 %

Evaluación Permanente 2 : 20 %

Examen Final : 30 %

Referencias

[Clote and Backofen, 2000] Clote, P. and Backofen, R. (2000). *Computational Molecular Biology, An Introduction*. Wiley.

[Jones and Pevzner, 2004] Jones, N. A. and Pevzner, P. A. (2004). *An Introduction to Bioinformatics Algorithms*. The MIT Press.

[Mount, 2001] Mount, D. W. (2001). *Bioinformatics: Sequence and Genome Analysis*. Cold Spring Harbor Laboratory Press.