

Cours IFT3295/BIN6000/IFT6291 Bio-Informatique

1 Description

Ce cours est destiné aux étudiants désireux de découvrir le domaine de la biologie computationnelle. Aucun prérequis en biologie n'est nécessaire: le cours commencera par une introduction au code génétique. Des connaissances biologiques supplémentaires seront introduites, au besoin, tout au long de la session. Cependant, des bases suffisantes en algorithmique et programmation sont nécessaires.

Les thèmes biologiques abordés concernent l'analyse des séquences biologiques (ADN, ARN, protéines) et l'étude de leur évolution. Le cours sera orienté vers les **méthodes algorithmiques déterministes** utilisées pour aborder les problématiques biologiques soulevées.

2 Objectifs

L'objectif de ce cours est d'**initier l'étudiant au domaine de la biologie computationnelle**, qui consiste à formaliser un problème biologique, et à développer les outils algorithmiques et mathématiques permettant de le résoudre. À l'issue de ce cours, l'étudiant sera **familiarisé avec les problématiques bio-informatiques classiques liées à l'analyse et au traitement des séquences biologiques, ainsi qu'avec les méthodes algorithmiques les plus utilisées pour les résoudre**.

Dans le cadre du cours BIN6000/IFT6295, l'objectif est également d'**initier l'étudiant à la recherche en biologie computationnelle**. Dans ce cadre, l'étudiant devra approfondir un ou plusieurs articles de recherche contenant une méthodologie algorithmique originale et en faire une présentation orale.

3 Programme

Les heures indiquées sont des estimations d'heures de cours. Les heures manquantes correspondent aux heures de démonstration, et aux examens.

1. Introduction au cours et à la Bio-Informatique (**3h**)
2. Alignement de séquences biologiques; méthodes exactes (**6h**)
3. Prédiction de structures secondaires d'ARN (**3h**)
4. Alignement multiple (**4h**)

5. Recherche de motifs: méthodes exactes et heuristiques **(7h)**
6. Arbres des suffixes **(4h)**
7. Phylogénie: Modèles évolutifs, inférence **(8h)**

4 Quelques manuels de référence

- *Algorithms on Strings, Trees and Sequences - Computer Science and Computational Biology*, Dan Gusfield, Cambridge University Press, 1997.
- *Concepts et méthodes en phylogénie moléculaire*, Guy Perrière et Céline Brochier-Armanet, Springer, 2010.
- *Handbook of Computational Molecular Biology*, Srinivas Aluri ed., Chapman & Hall/CRC Computer and Information Science Series, 2006.
- *An Introduction to Bioinformatics Algorithms*, Neil C. Jones and Pavel A. Pevzner, The MIT Press, 2004.
- *Bioinformatics Algorithms - Techniques and Applications*, Ion I. Mandoiu and Alexander Zelikovsky eds., Wiley, 2008.
- *Algorithms in Computational Molecular Biology - Techniques, Approaches and Applications*, Mourad Elloumi and Albert Y. Zomaya eds., Wiley Series in Bioinformatics, 2011.
- *Introduction to Computational Molecular Biology*, J. Setubal and J. Meidanis, PWS Publishing Company, 1997.
- *Biological sequence analysis - Probabilistic models of proteins and nucleic acids*, R. Durbin, S. Eddy, A. Krogh and G. Mitchison, 1998.

5 Mode d'évaluation

Notez qu'il y a un seuil de 50% de moyenne Intra+Examen comme condition de réussite.

5.1 IFT3295

1. Examen Intra 30%
2. Examen Final 40%
3. TP 30%

5.2 BIN6000/IFT6291

1. Examen Intra 30%
2. Examen Final 40%
3. Présentation d'articles 30%

L'étudiant doit identifier et valider l'article ou le sujet de recherche avec la professeure avant le 17 octobre.

L'article choisi devra être en lien direct avec le cours et devra être orienté vers une méthodologie algorithmique originale et innovante.

Le travail final constituera en une présentation orale de 35 minutes dont le temps sera réparti de la façon suivante:

- Un maximum de 10mn de présentation pédagogique d'un concept informatique ou biologique lié à l'article et non présenté en cours;
- Présentation de l'article, avec une attention particulière portée à la méthodologie informatique et algorithmique considérée;
- 10 minutes de période de questions.

La notation se basera sur les critères suivants:

- Présentation orale (25%): Compréhension de la méthodologie et clarté dans les explications et la présentation de l'algorithme et de sa complexité en temps et en espace. Qualité générale de la présentation. Réponse aux questions.
- Participation aux périodes de question (5%)

6 Règlement sur le plagiat

Le plagiat est passible de sanctions allant jusqu'à l'exclusion du programme.

- Voir l'intégralité du règlement à : <http://integrite.umontreal.ca>.
- **Extrait:** Constituent notamment un plagiat ou une fraude:
 1. le copiage;
 2. la substitution de personne lors d'une Évaluation;
 3. l'exécution par une autre personne d'une Évaluation;
 4. l'utilisation totale ou partielle, littérale ou déguisée, d'un texte, d'un tableau, d'une image, d'un exposé, d'un enregistrement ou de toute autre création d'autrui, publié ou non, sans indication de référence adéquate à l'occasion d'une Évaluation;
 5. l'utilisation d'une traduction totale ou partielle d'un texte d'autrui, publié ou non, sans indication de référence adéquate à l'occasion d'une Évaluation;
 6. la reproduction, la distribution, et l'obtention par vol, manœuvre, corruption ou par tout autre moyen, de questions ou de réponses d'examen ou de tout autre document non autorisé;
 7. la sollicitation, l'offre ou l'échange d'information pendant une Évaluation.