

Université de
Sherbrooke

Département d'informatique

BIN 702 – Algorithmes pour la bio-informatique

Plan d'activité pédagogique

Hiver 2021

Enseignant

Manuel Lafond

Courriel : manuel.lafond@usherbrooke.ca

Local : D4-2010

Téléphone : +1 819 821-8000 x62034

Disponibilités : Mardi de 13h30 à 17h00

Responsable(s) : Direction du département

Site web du cours : <http://info.usherbrooke.ca/mlafond/BIN702/>

Horaire

Exposé magistral :	Lundi	13h30 à 15h20	salle Distance
	Mercredi	8h00 à 9h50	salle Distance

Description officielle de l'activité pédagogique¹

Cibles de formation :	Apprendre les algorithmes et techniques appliqués aux structures en biologie moléculaire (arbres de phylogénie, structures tridimensionnelles, agglomérats moléculaires). Acquérir des habiletés pour développer des applications pour la bio-informatique.
Contenu :	Comparaison et alignement des séquences biologiques. Recherche de motifs. Alignement multiple. Prédiction de la structure secondaire et tertiaire de l'ARN et algorithmes de repliement. Comparaison des structures secondaires de l'ARN. Structures des protéines (secondaire, tertiaire, quaternaire). Classification des protéines.
Crédits	3
Organisation	3 heures d'exposé magistral par semaine 6 heures de travail personnel par semaine
Particularités	Aucune

¹<https://www.usherbrooke.ca/admission/fiches-cours/bin702>

1 Présentation

Cette section présente les cibles de formation spécifiques et le contenu détaillé de l'activité pédagogique. Cette section, non modifiable sans l'approbation du comité de programme du Département d'informatique, constitue la version officielle.

1.1 Mise en contexte

Le cours BIN702 - Algorithmes pour la bio-informatique est un cours à option de 1^{er}, 2^e et 3^e cycles en informatique. Nous y introduisons les principales structures de données, algorithmes et techniques utilisés dans le domaine de la biologie computationnelle. Il ne présuppose aucune connaissance en biologie. Les connaissances en sciences de la vie nécessaires seront introduites au fur et à mesure. Des bases en structures de données sont requises, et des connaissances minimales en algorithmique sont souhaitables.

1.2 Cibles de formation spécifiques

Ce cours vise des étudiantes et des étudiants qui souhaitent découvrir la biologie computationnelle ou approfondir leur connaissance des modèles et algorithmes sur lesquels reposent les outils bio-informatiques. À la fin de cette activité pédagogique, l'étudiante ou l'étudiant se sera capable

1. de comprendre et de concevoir diverses modélisations de problèmes en biologie computationnelle ;
2. de développer des algorithmes pour ces problèmes et de les comparer en analysant leur complexité et en les analysant expérimentalement ;
3. de se familiariser avec les structures de données et les méthodes principales de comparaison, de recherche et d'extraction d'information, de prédiction de structures de séquences biologiques ;
4. de connaître les modèles et algorithmes établis de reconstruction d'histoires évolutives de séquences et structures biologiques.

1.3 Contenu détaillé

Thème	Contenu	Nbr. d'heures	Objectifs
1	Introduction à la bio-informatique	3	1
2	Alignement de séquences biologiques : approches exactes et heuristique BLAST	6	1, 2, 3
3	Séquençage de génomes (avec ou sans référence)	3	1, 2, 3
4	Recherche de motifs, filtrage de chaînes et arbres de suffixe	5	1, 2, 3
5	Modèles de Markov cachés et modélisation de séquences	4	1, 2, 3
6	Alignement multiple de séquences	5	1, 2, 3, 4
7	Reconstruction phylogénétique pour l'inférence d'histoires évolutives	6	1, 2, 4
8	Prédiction et recherche de structures d'ARN	3	1, 2, 3, 4
9	Réarrangements génomiques	3	1, 2, 4

2 Organisation

Cette section propre à l'approche pédagogique de chaque enseignante ou enseignant présente la méthode pédagogique, le calendrier, le barème et la procédure d'évaluation ainsi que l'échéancier des travaux. Cette section doit être cohérente avec le contenu de la section précédente.

2.1 Méthode pédagogique

Chaque semaine, il y aura trois (3) heures de cours magistraux. Des exercices optionnels seront fournis aux étudiant(e)s afin de mettre en pratique les concepts théoriques présentés et certains de ces exercices seront discutés en classe lorsque le temps le permettra. Les étudiant(e)s seront amené(e)s à développer leurs compétences en bio-informatique à l'aide d'un projet de session de nature théorique ou pratique, au choix. Les étudiant(e)s du 2e et 3e cycles devront consolider les concepts vus en cours par la présentation/vulgarisation/critique d'articles scientifiques liés au sujet du cours (voir les informations sur les évaluations).

Compte tenu du contexte actuel (pandémie due au COVID-19), il se peut que le cours ait lieu en totalité ou en partie à distance d'une façon différente de ce qui est énoncé ci-dessus. Notez que vous en serez informés rapidement si tel est le cas.

2.2 Calendrier

Semaine	Date	Thème
1	2021-01-11	1
2	2021-01-18	2
3	2021-01-25	2
4	2021-02-01	4
5	2021-02-08	4 et 6
6	2021-02-15	6
7	2021-02-22	Examen périodique
8	2021-03-01	Relâche
9	2021-03-08	7
10	2021-03-15	3 et 7
11	2021-03-22	3 et 8
12	2021-03-29	5 et 8
13	2021-04-05	5 et 9
14	2021-04-12	Examen final
15	2021-04-19	Examen final

2.3 Évaluation

Projet de session	34 %
Examen intra	33 %
Examen final	33 %

Les examens se font à distance et doivent être remis de façon électronique. Toute forme de documentation est permise.

Chaque étudiant et étudiante devra entreprendre un projet de session. Le projet peut être réalisé de façon individuelle ou en équipe de deux. Les attentes seront plus élevées pour les projets en équipe. Le projet peut être de nature pratique (par exemple, implémenter un algorithme pour la bio-informatique et l'évaluer sur des données réelles et/ou simulées) ou théorique (par exemple, essayer de résoudre un problème ouvert ou d'améliorer la complexité d'un algorithme connu).

Pour les projets pratiques, le choix du langage de programmation est laissé au choix de l'étudiant(e). Pour les projets théoriques, il n'est pas requis de résoudre le problème étudié, mais il est important de démontrer que suffisamment de pistes ont été explorées et de discuter pourquoi elles n'ont pas réussi. La nature du projet ainsi que les dates et pondérations des livrables requis au cours de la session seront à discuter avec le professeur. Une banque de projets sera proposée, mais les étudiant(e)s sont encouragés à faire leurs propres propositions.

Aux 2e et 3e cycles, la compréhension d'un article de recherche en lien avec le cours est attendue dans le cadre du projet. Ceci peut prendre la forme d'un livrable séparé qui consiste en une présentation, en un texte de vulgarisation, ou autre, à discuter avec le professeur.

2.3.1 Qualité de la langue et de la présentation

Conformément à l'article 17 du règlement facultaire d'évaluation des apprentissages² l'enseignante ou l'enseignant peut retourner à l'étudiante ou à l'étudiant tout travail non conforme aux exigences quant à la qualité de la langue et aux normes de présentation.

2.3.2 Plagiat

Le plagiat consiste à utiliser des résultats obtenus par d'autres personnes afin de les faire passer pour sien et dans le dessein de tromper l'enseignante ou l'enseignant. Vous trouverez en annexe un document d'information relatif à l'intégrité intellectuelle qui fait état de l'article 9.4.1 du Règlement des études³. Lors de la correction de tout travail individuel ou de groupe une attention spéciale sera portée au plagiat. Si une preuve de plagiat est attestée, elle sera traitée en conformité, entre autres, avec l'article 9.4.1 du Règlement des études de l'Université de Sherbrooke. L'étudiante ou l'étudiant peut s'exposer à de graves sanctions qui peuvent être soit l'attribution de la note E ou de la note zéro (0) pour un travail, un examen ou une activité évaluée, soit de reprendre un travail, un examen ou une activité pédagogique. Tout travail suspecté de plagiat sera transmis au Secrétaire de la Faculté des sciences. Ceci n'indique pas que vous n'avez pas le droit de coopérer entre deux équipes, tant que la rédaction finale des documents et la création du programme restent le fait de votre équipe. En cas de doute de plagiat, l'enseignante ou l'enseignant peut demander à l'équipe d'expliquer les notions ou le fonctionnement du code qu'elle ou qu'il considère comme étant plagié. En cas d'incertitude, ne pas hésiter à demander conseil et assistance à l'enseignante ou l'enseignant afin d'éviter toute situation délicate par la suite.

2.4 Échéancier des travaux

Projet de session	Sujet	Réception	Remise	Points
Projet de session	Projet de session	À définir	À définir	34

2.5 Utilisation d'appareils électroniques et du courriel

Selon le règlement complémentaire des études, section 4.2.3⁴, l'utilisation d'ordinateurs, de cellulaires ou de tablettes pendant une prestation est interdite à condition que leur usage soit explicitement permise dans le plan de cours.

Dans ce cours, l'usage de téléphones cellulaires, de tablettes ou d'ordinateurs est autorisées. Cette permission peut être retirée en tout temps si leur usage entraîne des abus.

Tel qu'indiqué dans le règlement universitaire des études, section 4.2.3⁵, toute utilisation d'appareils de captation de la voix ou de l'image exige la permission de la personne enseignante.

Note : L'utilisation du courriel est recommandée pour poser vos questions.

3 Matériel nécessaire pour l'activité pédagogique

Le contenu du cours est largement basé sur le livre [7]. Il n'est pas obligatoire, mais fortement recommandé.

²https://www.usherbrooke.ca/sciences/fileadmin/sites/sciences/Etudiants_actuels/Informations_academiques_et_reglements/2017-10-27_Reglement_facultaire_-_evaluation_des_apprentissages.pdf

³<https://www.usherbrooke.ca/registraire/droits-et-responsabilites/reglement-des-etudes/>

⁴https://www.usherbrooke.ca/sciences/fileadmin/sites/sciences/documents/Intranet/Informations_academiques/Sciences_Reglement_complementaire_2017-05-09.pdf

⁵<https://www.usherbrooke.ca/registraire/droits-et-responsabilites/reglement-des-etudes/>

4 Références

- [1] DAN GUSFIELD : *Algorithms on Strings, Trees and Sequences - Computer Science and Computational Biology*. Cambridge University Press, 1997.
- [2] JOSEPH FELSENSTEIN : *Inferring phylogenies*. Sinauer Associates Inc, 2002.
- [3] JOÃO SETUBAL AND JOÃO MEIDANIS : *Introduction to Computational Molecular Biology*. PWS Publishing Company, 1997.
- [4] NEIL C. JONES AND PAVEL A. PEVZNER : *An Introduction to Bioinformatics Algorithms*. The MIT Press, 2004.
- [5] OLIVIER GASCUEL : *Mathematics of evolution & phylogeny*. Oxford University Press, 2004.
- [6] RICHARD DURBIN, SEAN R. EDDY, ANDERS KROGH AND GRAEME MITCHISON : *Biological sequence analysis - Probabilistic models of proteins and nucleic acids*. Cambridge University Press, 1998.
- [7] WING-KING SUNG : *Algorithms in Bioinformatics - A practical introduction*. CRC Mathematical and Computational Biology, 2009.

L'intégrité intellectuelle passe, notamment, par la reconnaissance des sources utilisées. À l'Université de Sherbrooke, on y veille!

Extrait du Règlement des études (Règlement 2575-009)

9.4.1 DÉLITS RELATIFS AUX ÉTUDES

Un délit relatif aux études désigne tout acte trompeur ou toute tentative de commettre un tel acte, quant au rendement scolaire ou une exigence relative à une activité pédagogique, à un programme ou à un parcours libre.

Sont notamment considérés comme un délit relatif aux études les faits suivants :

- a) commettre un plagiat, soit faire passer ou tenter de faire passer pour sien, dans une production évaluée, le travail d'une autre personne ou des passages ou des idées tirés de l'œuvre d'autrui (ce qui inclut notamment le fait de ne pas indiquer la source d'une production, d'un passage ou d'une idée tirée de l'œuvre d'autrui);
 - b) commettre un autoplagiat, soit soumettre, sans autorisation préalable, une même production, en tout ou en partie, à plus d'une activité pédagogique ou dans une même activité pédagogique (notamment en cas de reprise);
 - c) usurper l'identité d'une autre personne ou procéder à une substitution de personne lors d'une production évaluée ou de toute autre prestation obligatoire;
 - d) fournir ou obtenir toute aide non autorisée, qu'elle soit collective ou individuelle, pour une production faisant l'objet d'une évaluation;
 - e) obtenir par vol ou toute autre manœuvre frauduleuse, posséder ou utiliser du matériel de toute forme (incluant le numérique) non autorisé avant ou pendant une production faisant l'objet d'une évaluation;
 - f) copier, contrefaire ou falsifier un document pour l'évaluation d'une activité pédagogique;
- [...]

Par plagiat, on entend notamment :

- Copier intégralement une phrase ou un passage d'un livre, d'un article de journal ou de revue, d'une page Web ou de tout autre document en omettant d'en mentionner la source ou de le mettre entre guillemets;
- reproduire des présentations, des dessins, des photographies, des graphiques, des données... sans en préciser la provenance et, dans certains cas, sans en avoir obtenu la permission de reproduire;
- utiliser, en tout ou en partie, du matériel sonore, graphique ou visuel, des pages Internet, du code de programme informatique ou des éléments de logiciel, des données ou résultats d'expérimentation ou toute autre information en provenance d'autrui en le faisant passer pour sien ou sans en citer les sources;
- résumer ou paraphraser l'idée d'un auteur sans en indiquer la source;
- traduire en partie ou en totalité un texte en omettant d'en mentionner la source ou de le mettre entre guillemets ;
- utiliser le travail d'un autre et le présenter comme sien (et ce, même si cette personne a donné son accord);
- acheter un travail sur le Web ou ailleurs et le faire passer pour sien;
- utiliser sans autorisation le même travail pour deux activités différentes (autoplagiat).

Autrement dit : mentionnez vos sources
