McGill

AC-04-18
New Course Proposal Form

(09/2003)

1. Will this new course affect a current program? ☒ Yes ☐ No
If "yes", has a Program Revision Form been submitted concurrently? ☐ Yes ☒ No

2. Teaching Department: School of Computer Science

3. Administering Faculty/Unit: Faculty of Science

4. Campus (Downtown, Macdonald, Off Campus, Distance Ed, Other – specify)
Downtown

5. Effective Term of Implementation (Ex. Sept. 2004 = 200409)
Term: 200501

6. Course Title (Limit 30 Characters) - required for all courses:
Mining biological sequences

8. Course Title to Appear in the Calendar (optional) (Limit 59 characters):
Note: This can ONLY be an expansion of word(s) abbreviated in the 30 character course title above.
Algorithms for mining biological sequences

9. Credit Weight (or CEU's for non-credit CE courses):
4 credits

7. Course Number(s)
Indicate course number & the number of terms spanned: (tick all that apply)

Subject/course number: COMP 680

Course(s) Span:
☒ **1 term**
☐ **2 consecutive terms (D1, D2)**
☐ **2 non-consecutive terms (N1, N2)**
☐ **3 terms (J1, J2, J3)**

10. Schedule Type(s):
(Enter all that apply – see form, STVSCHD in Banner for a complete list.)
(i.e. Lecture, Labs, Tutorial)

	Hours per Week		Hours per Week		Hours per Week
Lecture	3	Project	3		

Total Hours per Week: 6

This course will involve a substantial final project, and students will also have to prepare a 60-minute in-class paper presentation. This amount of work is beyond the usual homework and assignment workload.

Total Number of Weeks: 13

11. Projected Enrolment:
25

12. Prerequisite(s) (Courses or Tests)
Specify course number(s) or name(s) of test(s):

COMP 462 or equivalent

If the student does not have a prerequisite should web registration be blocked?
☐ **Yes** ☒ **No**

If "Yes" complete A and B:

A. Indicate minimum grade or test score(s) the student must attain in prerequisite course(s) or test(s):

B. Can the prerequisite course(s) or test(s) be taken in the same term as this course?
☐ **Yes** ☒ **No**

13. Corequisite(s) Course Number(s):
Specify course number(s) and title(s):

If the student does not register for the corequisite in the same term should web registration be blocked?
☐ **Yes** ☐ **No**

14. Consultation Reports Attached
☐ **Yes** ☒ **N/A**

15. Additional Course Charges (must be approved by the Fee Policy Committee)

Description of Fee (e.g. screening fee)	**Amount**

16. Requires Teaching, Physical, or Financial Resources Not Currently Available (attach explanation)
☐ **Yes** ☒ **No**

17. Other Information (specify):

18. Course Description
(as it will appear in the Calendar [maximum 50 words]):
(N.B. Faculty of Medicine must append complete course outline)

Advanced algorithms for the annotation of biological sequences. Algorithms and heuristics for pair-wise and multiple sequence alignment. Gene-finding with hidden Markov models and their variants. Motifs discovery techniques: over-representation and phylogenetic footprinting approaches. RNA secondary structure prediction and modeling. Detection and classification of repetitive elements. Representation and annotation of protein domains.

19. Supplementary information to appear in the Calendar in addition to the course description.
Such as: registration restriction(s), prerequisite(s), corequisite(s), equivalent course(s), contact hours, enrolment limitations, language of instruction etc.
Please enter the information as it should appear in the calendar notes.

COMP 462 or with instructor's permission

20. Rationale

The biotechnology revolution relies heavily on the ability to analyze the huge amount of data produced. Biological sequences (DNA and proteins) are of little use without a functional annotation, which can best be obtained through computational analysis (bioinformatics). This course will explore the cutting-edge algorithmic and machine learning techniques required for mining biological sequences, thus preparing students for innovative work in the industry and in research institutions while helping fuel research in applied molecular biology.
With the move of COMP 562 (Computational biology methods) to COMP 462 (Computational biology methods), the School of Computer Science does not offer a graduate course focused on algorithms for the analysis of biological sequences. The course proposed here gives the opportunity to graduate students to get a solid background on sequence analysis, focusing on algorithmic, statistical, and machine learning approaches. Others (sequence alignment and hidden Markov models) will have already been introduced in COMP 462. These topics will only be very briefly reviewed before taking over where COMP 462 left off, introducing advanced algorithms and statistics not covered by COMP 462. The course is complementary to COMP 563 (Molecular Evolution Theory) and COMP 564 (Computation Gene Regulation) but shares little content with them. Because of its computational focus, the course proposed has no significant overlap with any other course offered at McGill. This course will involve an important final project, and students will also have to prepare a 60-minute in-class paper presentation. This amount of work is beyond the usual assignment workload.

INFORMATION FOR ADMISSIONS, RECRUITMENT & REGISTRAR'S OFFICE

To be completed by the Faculty	*To be completed by ARR*	*For Continuing Education Use*
Slot Course: ☐ **Yes** ☐ **No**	**CIP Code**	**CE Admin. Unit :**
Thesis Component: ☐ **Yes** ☐ **No**		**CE Non-Grant Courses:**
		Flat Rate: CdnFlat Rate: ☐ **Yes** ☐ **N/A**

21. Approvals:

Routing Sequence	**Departmental Meeting**	**Departmental Chair**	**Other Faculty**	**Curric/Academic Committee**	**Faculty**	**SCTP**
Name		Denis Thérien				
Signature						
Date		17 Sept. 2004				

Departmental Contact Person (name/phone/email)	