

POST-DOC BIOINFORMATIQUE ET APPRENTISSAGE – LYON : CDD 18 MOIS

La bioinformatique constitue un axe de recherche multidisciplinaire à la frontière de plusieurs disciplines scientifiques : la biologie, l'informatique, la pharmacologie, etc ... A l'ère de la postgénomique, un des challenges de la bioinformatique est la prédiction de la structure 3D d'une protéine à partir de sa séquence (enchaînement d'acide aminé).

Le projet GENOTO3D, soutenu par l'ACI « Masse de Donnée » va notamment porter ses efforts sur le développement d'approches prédictives d'interaction 3D d'acides aminés éloignés dans la séquence. Nous poursuivons ainsi des recherches que nous avons initiées dans le cadre du groupe de travail « Apprentissage et séquences » de l'AS STIC « Apprentissage et bioinformatique ».

Ce projet animé par Yann Guermeur (LORIA, Nancy) est un consortium de plusieurs équipes issues de la biologie et de l'informatique :

- Pôle de BioInformatique Lyonnais (PBIL-IBCP, UMR 5086, <http://pbil.ibcp.fr>), Lyon
- Modèles informatiques en biologie moléculaire (MODBIO), LORIA – UMR 7503, Nancy
- Laboratoire d'Informatique Fondamentale (LIF), UMR 6166, Marseille
- Systèmes et modèles biologiques, bioinformatique et séquences (Symbiose), IRISA / INRIA Rennes
- Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (LIRMM), UMR5506, Montpellier
- Mathématique Informatique et Génome, INRA, Jouy en Josas

Dans ce cadre la, notre laboratoire bénéficie d'un financement pour embaucher un post-doc informaticien pour une durée de 18 mois. Des compétences fines dans certains des domaines suivants sont souhaitées : apprentissage automatique, probabilités, statistique, programmation mathématique et combinatoire. Une sensibilité aux problèmes biologiques est aussi la bienvenue.

Les domaines de l'apprentissage que nous explorerons de manière privilégiée sont l'apprentissage statistique, à travers la conception et la mise en œuvre de machines à noyau et de réseaux de neurones récurrents, ainsi que la théorie des automates, avec en particulier l'inférence de machines de reconnaissance probabilistes (automates probabilistes, HMMs,...). Leur mise en œuvre combinée devra en particulier permettre de prendre compte les dépendances à long terme, encore mal exploitées par les méthodes de prédiction classiques.

Les retombées attendues sont importantes, en effet des méthodes fines de prédiction de la structure tertiaire des protéines permettront une meilleure connaissance «mécanistique» des protéines. Ceci intéresse donc l'ensemble de la communauté biologique aussi bien académique qu'industrielle (en particulier l'industrie pharmaceutique). Du point de vue informatique, il est clair que les méthodes développées devraient donner naissance à des outils largement utiles à des domaines de traitement intensif de séquences comme le traitement de textes en langage naturel. Le domaine de l'apprentissage automatique a toujours fortement bénéficié de l'analyse systématique et approfondie d'applications spécifiques.

Merci d'adresser un CV ainsi qu'un résumé de vos travaux à :

Christophe Geourjon

Laboratoire de Bioinformatique et RMN structurales

Pôle de BioInformatique Lyonnais - PBIL-IBCP

7, Passage du Vercors - 69 367 Lyon cedex 07

Web : <http://pbil.ibcp.fr> - Mail : c.geourjon@ibcp.fr - Tel : 04 72 72 26 47