

Master Recherche M2

Sujet de stage : Etude préliminaire à la conception d'une ontologie et d'un moteur d'inférence pour la biologie systémique

Encadrants :

Christine Froidevaux (Professeur), chris AT lri.fr

Anne Poupon (CR1, INRA Tours), anne.poupon AT tours.inra.fr

Lieu du stage :

Equipe Bioinformatique, Laboratoire de Recherche en Informatique, Université Paris-Sud, Orsay, France

Prérequis :

Aucune connaissance en biologie n'est prérequisée, mais une certaine curiosité pour la biologie est attendue.

Contexte :

L'étude d'un système biologique débute par une accumulation de données, souvent obtenues par des méthodes expérimentales à haut-débit, suivie de la construction manuelle d'un modèle statique par l'expert biologiste, et enfin une paramétrisation et une simulation du modèle dynamique par des méthodes informatiques et mathématiques spécialisées. L'étape intermédiaire de ce processus, à savoir la conception du modèle statique, constitue un goulot d'étranglement. L'expert réalise souvent des choix quant aux données prises en compte, qui rendent le modèle biaisé et incomplet. Pour rendre performante cette étape, deux types de développements sont nécessaires : (i) un système de gestion des données expérimentales et non expérimentales, permettant de mettre à disposition l'ensemble des données nécessaires à la modélisation ; et (ii) une méthode permettant d'inférer le modèle statique directement à partir ces données. L'un et l'autre de ces développements nécessitent la création d'une ontologie rassemblant l'ensemble des concepts du domaine. De plus, l'implémentation d'outils permettant l'application pratique de ces méthodes nécessite d'être capables de représenter l'ontologie dans le langage UML.

Travail :

La réalisation d'une méthode permettant d'inférer de manière automatique un réseau biologique à partir des données expérimentales est un projet qui s'étalera sur plusieurs années. L'objet du stage est la préparation des méthodes et des outils nécessaires.

Pour cela, nous construirons une ontologie simplifiée correspondant à un sous-domaine cohérent. Cette ontologie réduite devra en particulier permettre de rendre compte de la complexité des règles de l'ontologie complète. Il s'agira de (i) déterminer un formalisme adéquat (tel que OWL) pour représenter l'ontologie et permettre de raisonner efficacement ; (ii) déterminer un moteur d'inférence adéquat (tel que F-OWL) ; et (iii) concevoir un système de conversion entre le formalisme choisi pour l'ontologie d'une part, et UML d'autre part.