

*Proposition de stage INRA*

## **Comparaison de bibliothèques pour l'inférence dans un réseau bayésien dynamique en Matlab**

### **Contexte**

L'une des principales méthodes employées par les agriculteurs pour lutter contre les bioagresseurs des cultures (par exemples les mauvaises herbes) est l'utilisation de pesticides. Ces produits sont une solution radicale mais ils constituent une menace pour l'environnement et la santé. On s'intéresse à une meilleure utilisation de l'information mise à disposition par des réseaux d'épidémio-surveillance pour limiter les traitements de culture.

Pour ce faire, un modèle simplifié de gestion des bioagresseurs a été développé sous forme de Réseau Bayésien Dynamique (RBD). Un RBD est un modèle stochastique qui étend la notion de réseau bayésien. A la différence d'un réseau bayésien, un RBD permet de représenter l'évolution des variables aléatoires sur des pas de temps discrets. Le modèle développé formalise à la fois (i) la propagation dans le temps d'une épidémie d'une parcelle à l'autre, (ii) l'observation de l'épidémie sur certaines parcelles à chaque période de temps (information fournie par les réseaux d'épidémio-surveillance), (iii) la prise de décision de traiter ou non les parcelles infectées.

Ce modèle développé en Matlab va être utilisé pour évaluer l'incidence de différentes stratégies d'échantillonnage des observations (patterns d'observation) sur le nombre de traitements et l'état d'infection des parcelles. Depuis un état initial donné, l'évolution du système est modélisé sur  $T$  périodes de temps. La simulation demande de calculer la probabilité d'infestation d'une parcelle sachant les observations. C'est un calcul d'inférence pour lequel plusieurs algorithmes ont été développé.

### **Sujet**

L'objectif du stage est de tester différentes fonctions d'inférence appelables depuis Matlab pour évaluer la probabilité d'infection d'une parcelle.

Dans un premier temps, il conviendra de sélectionner les fonctions considérées parmi les bibliothèques possibles : BNT (<https://code.google.com/archive/p/bnt>), PMTK (<https://github.com/probml/pmtk3>), OpenGM2 (<http://hci.iwr.uni-heidelberg.de/opengm2>), libDAI (<https://staff.fnwi.uva.nl/j.m.mooij/libDAI>) et dimple (<http://dimple.problog.org>).

Dans un deuxième temps, il conviendra de recoder le modèle dans les différentes modélisations (réseau bayésien, factor graph) correspondantes aux algorithmes sélectionnés. Pour quelques configurations (parcellaire, état initial, historique et lieux des observations) définies une comparaison des algorithmes sera faite en considérant les critères temps d'exécution, validité du résultat et facilité d'appel.

### **Compétences recherchées**

Des connaissances en mathématique (probabilité) et en programmation Matlab sont nécessaires.

### **Durée**

Deux à trois mois (entre avril et septembre 2016).

### **Contacteur**

Marie-Josée Cros

Unité de Mathématiques et Informatique Appliquées de Toulouse

INRA, BP 52627, 31326 Castanet cedex

Tel : 05-61-28-52-84

Email : [mjcros@toulouse.inra.fr](mailto:mjcros@toulouse.inra.fr)