



**Département de Génie Informatique  
et Génie Logiciel**

**INF8225**

**I.A. : techniques probabilistes et d'apprentissage**

**Laboratoire 1**

**TP 1**

## Objectif du laboratoire

Ce laboratoire est une introduction au calcul de probabilité de distributions jointes dans un réseau de Bayes. L'objectif du laboratoire est de permettre à l'étudiant de se familiariser avec les logiciels *Netica* et *MATLAB* accompagnés de la boîte à outils *Pmtk3*.

## Description des tâches

### Familiarisation avec Netica

Le logiciel permet de construire et maintenir rapidement un réseau de Bayes. Il y est facile de tester l'impact global en termes de probabilité des nœuds du réseau selon un événement qui modifie la distribution. Rendez-vous à l'adresse suivante pour télécharger le logiciel Netica (vous utiliserez la version gratuite limitée):

<http://www.norsys.com/download.html>

Vous devrez installer (dézipper) le logiciel dans votre espace personnel (Y:/). Vous trouverez sur le site un bon tutoriel pour comprendre le logiciel. Les deux premières sections (A et B) sont à lire et à comprendre :

[http://www.norsys.com/tutorials/netica/nt\\_toc\\_A.htm](http://www.norsys.com/tutorials/netica/nt_toc_A.htm)

### Familiarisation avec MATLAB

Si vous n'avez jamais touché à MATLAB, il est fortement recommandé de vous familiariser avec le logiciel en suivant les tutoriels officiels :

[http://www.mathworks.com/academia/student\\_center/tutorials/launchpad.html](http://www.mathworks.com/academia/student_center/tutorials/launchpad.html)

### Préparer l'environnement de développement MATLAB

Téléchargez la boîte d'outils *pmtk3* développée par une équipe de UBC. Elle contient plusieurs fonctionnalités désirables qui vous faciliteront le travail lors des laboratoires.

<http://code.google.com/p/pmtk3/> ou <https://github.com/probml/pmtk3>

1. Dézippez le contenu du fichier téléchargé dans votre espace personnel attribué par l'école (ex : Y:\DOSSIER\_DE\_VOTRE\_CHOIX)
2. Les instructions suivantes (le test d'initialisation est facultatif) devront être faites chaque fois que vous voudrez utiliser la boîte d'outils *Pmtk3* dans MATLAB

Ouvrez MATLAB et entrez les commandes :

- `cd 'Y:\DOSSIER_DE_VOTRE_CHOIX'`
- `initPmtk3` (Soyez patient et attendez la fin de l'initialisation)
- `testPmtk3` (Soyez patient et attendez la fin du test d'installation)

Normalement, à la fin des tests, vous devriez obtenir le message : `All Tests Passed`  
`initPmtk3`

3. Vous devez maintenant étudier les deux exemples suivants faisant utilisation de la boîte à outils *Pmtk3* :

a) Sprinkler :

<http://pmtk3.googlecode.com/svn/trunk/docs/demoOutput/otherDemos/sprinklerGMdemo.html>

b) Lung Cancer :

<http://pmtk3.googlecode.com/svn/trunk/docs/demoOutput/otherDemos/lungcancerGMdemo.html>

Notez qu'il y a plusieurs autres exemples ici :

<http://pmtk3.googlecode.com/svn/trunk/docs/demoOutput/otherDemos/index.html>

## Question 1

Expliquer à l'aide d'exemples chacun des phénomènes suivants (les parties 3.2, 3.3 et 3.4 du tutoriel, utiliser un exemple différent) :

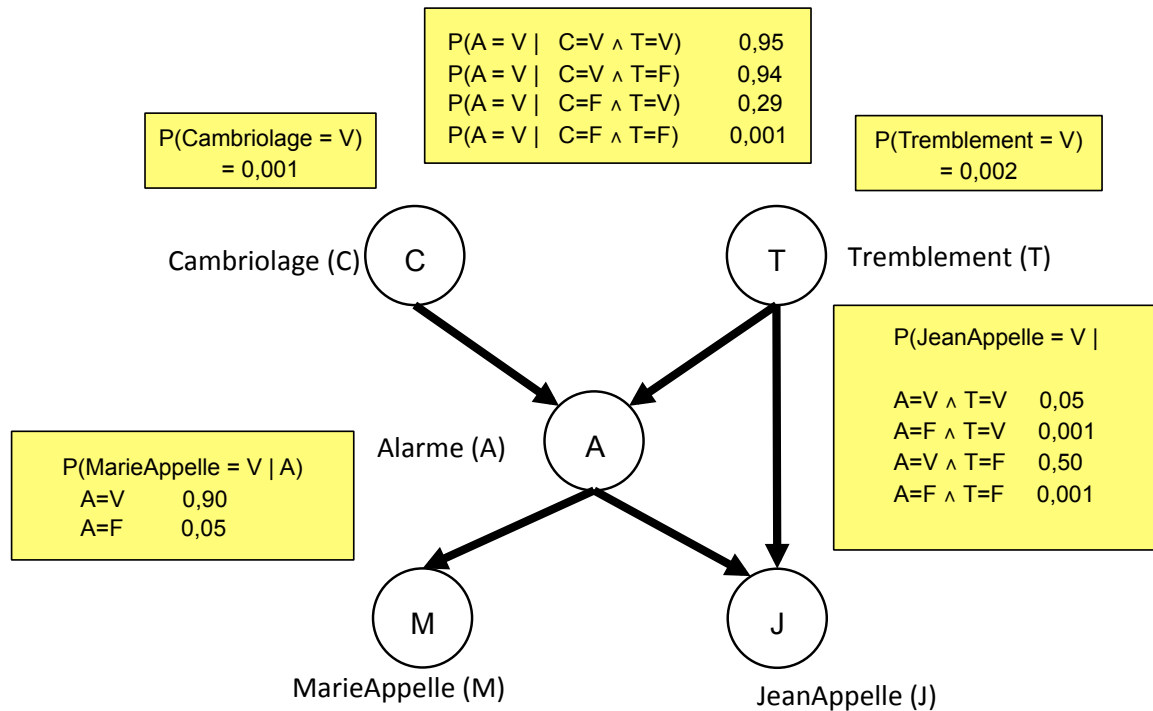
- (1) « explaining away »
- (2) « serial blocking »
- (3) « divergent blocking »

Expliquez clairement chaque phénomène en termes probabilistes, créez un réseau avec PMTK pour montrer chacun de ces effets et interprétez à haut niveau les observations.

**Barème : 12 pts (4 par phénomène : 2 points pour l'implémentation concrète d'exemple (il faut remettre le code), 1 pour l'explication théorique et 1 pour l'explication pratique).**

## Question 2

a) Créez un fichier « .m » « mkDetectorCambriolageDgm.m » pour l'exemple suivant :



Préparez un rapport avec les éléments similaires aux derniers deux exemples, spécifiquement :

b) Montrez le code et l'**histogramme pour la probabilité jointe** de votre modèle.

c) Utilisez PMTK pour calculer les marginales conditionnelles suivantes :

$P(\text{Cambriolage} = V \mid \text{MarieAppelle} = V, \text{JeanAppelle} = F)$ ,

$P(\text{Cambriolage} = V \mid \text{MarieAppelle} = F, \text{JeanAppelle} = V)$ ,

$P(\text{Cambriolage} = V \mid \text{MarieAppelle} = V, \text{JeanAppelle} = V)$ ,

$P(\text{Cambriolage} = V \mid \text{MarieAppelle} = F, \text{JeanAppelle} = F)$ ,

$P(\text{Cambriolage} = V \mid \text{MarieAppelle} = V)$ , (JeanAppelle=? ou la variable n'est pas observée)

$P(\text{Cambriolage} = V \mid \text{JeanAppelle} = V)$ , (MarieAppelle=? ou la variable n'est pas observée)

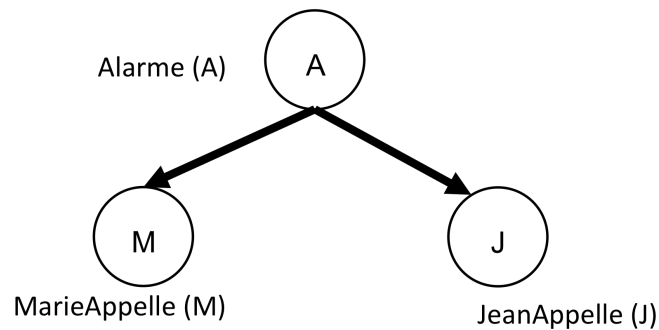
Ici encore, donnez le code et les résultats avec des valeurs numériques.

d) Calculez et montrez la probabilité marginale inconditionnelle pour chaque variable du modèle.

e) Ecrivez les equations pour expliquez comment à calculer :  $P(J)$  et  $P(C \mid J=V)$ .

### Question 3

a) Étant donné un ensemble de données avec  $N$  exemples et juste les observations pour les variables  $M$  et  $J$  (la variable  $A$  est toujours caché), donnez l'algorithme EM - avec les équations - pour estimer les paramètres d'un modèle comme illustré dans la figure suivante :



### Barème

#### Barème de correction Laboratoire 1, TP 1 (À remettre au début de notre prochain laboratoire)

Description des requis	Points alloués
Question 1	12
Question 2	13
a) 3 pts – fichier pour construire le réseau.	
b) 2 pts – 1 point pour le code, 1 point pour les valeurs correctes	
c) 4 pts – 2 points pour le code, 2 points pour les valeurs correctes	
d) 2 pts – 1 point pour le code, 1 point pour les valeurs correctes	
e) 2 pts – 1 point pour chaque équation correcte	
Question 3	5
<b>Total</b>	<b>30</b>