



UNIVERSITÉ
LIBRE
DE BRUXELLES

INFO-F311 - PROJET D'IA 2

RÉSEAUX BAYÉSIENS

Auteur:

Manuel ROCCA - 000596086

Professeurs:

Tom LENAERTS

Assistants:

Axel ABELS

Martin COLOT

Yannick MOLINGHEN

Pascal TRIBEL

Année académique 2025-2026

Contents

1	Introduction	2
2	Implémentation des expériences	2
2.1	Fonction de vraisemblance	2
2.2	Inférence par énumération	3
2.2.1	Disclaimer	3
2.2.2	Explication	3
3	Analyses de l'impact de paramètres sur la détection de gemmes	4
3.1	Impact du paramètre λ	4
4	Amélioration de la méthode d'inférence	5
5	Usage des LLMs	5
6	Conclusion	6

1 Introduction

Ce second projet d'Intelligence Artificielle a pour but d'appliquer la théorie concernant les réseaux Bayésiens. Cette structure de données repose fortement sur les probabilités, en particulier les probabilités conjointes. Donc, en plus de "bêtement" implémenter une structure de données, ce projet nous a permis de relier différents concepts provenants de divers cours du bachelier de sciences- informatiques (en particulier Probabilités et Statistiques, cours de BA2).

Voici une définition formelle du réseau de Bayes:

Définition 1 (Réseau de Bayes) *Un réseau bayésien est un modèle graphique probabiliste représentant un ensemble de variables aléatoires sous la forme d'un graphe orienté acyclique. Intuitivement, un réseau bayésien est à la fois :*

1. *un modèle de représentation des connaissances* ;
2. *une "machine à calculer" des probabilités conditionnelles* ;
3. *une base pour des systèmes d'aide à la décision.*¹

En bref, nous pouvons résumer à un réseau de Bayes à un *DAG* où chaque noeud possède une certaine probabilité dépendant des probabilités de chacun des noeuds parents.

2 Implémentation des expériences

Nous allons dans cette section détailler nos implémentations, comment nous avons procédé, à l'aide de quels outils et ou personnes.

2.1 Fonction de vraisemblance

Implémenter la fonction *likelihood()* a été la partie la plus triviale du projet. En effet, le mode opératoire est largement détaillé dans les consignes. Nous revenons tout de même dessus dans ce rapport par soucis de complétude.

Le but de cette fonction est de retourner une valeur indiquant la cohérence entre ce qui est observé (les mesures du sonar) et les hypothèses (position supposée des gemmes). Pour ce faire, nous commençons par calculer les observations, c'est-à-dire les distances euclidiennes entre les gemmes et la position actuelle. Nous calculons ensuite la distance d de Manhattan entre ce vecteur de distances observées et le vecteur de distances données.

Finalement nous appliquons la formule donnée dans les consignes:

$$e^{(-\frac{d}{\lambda})}$$

Le λ dans cette équation correspond à un hyperparamètre. Voici une définition pour se situer:

Définition 2 (Hyperparamètre) *Dans l'apprentissage automatique, un hyperparamètre est un paramètre dont la valeur est utilisée pour contrôler le processus d'apprentissage.*²

¹https://fr.wikipedia.org/wiki/Réseau_bayésien

²<https://fr.wikipedia.org/wiki/Hyperparamètre>

Un hyperparamètre est donc une variable fixée arbitrairement (ne dépend pas des données). Voyons donc ce que fait λ dans notre équation.

Nous avons donc une fonction exponentielle à exposant négatif. Si nous supposons d constant (ou que nous faisons abstraction de sa variation), il est facile de se rendre compte que plus λ augmente, plus la valeur de l'exposant sera petite et inversément. Plus la valeur totale de l'exposant tend vers 0, plus la valeur totale de l'expression approche de 1 et plus la valeur de l'exposant augmente, plus la valeur de l'expression tend vers 0.

Nous abordons tout ceci plus en détail dans la section 3.1

2.2 Inférence par énumération

2.2.1 Disclaimer

L'implémentation de l'inférence par énumération a été la partie la plus difficile de ce projet. En effet, elle est le cœur même de ce dernier.

Cette partie nous a causé bien des problèmes en termes de compréhension. C'est pourquoi, au terme d'une conversation avec un collègue étudiant (à savoir Romain Liefferinckx; son nom est cité dans le commit concerné), il nous a proposé son code. Forcément, récupérer un code et le copier n'a aucun intérêt académique. Nous avons donc veillé à bien comprendre le code fourni en utilisant principalement le cours mais également ChatGPT (qui a généré des commentaires pour mieux s'y retrouver).

En bref, pour prouver notre compréhension, nous allons expliquer au mieux l'implémentation dans la section qui suit.

2.2.2 Explication

Comme le demandent les consignes, il nous est demandé de ne mettre à jour que les noeuds de G . Pour ce faire, pour chaque noeud de G , nous allons appliquer le théorème de Bayes:

$$P(G | D) = \alpha P(D | G) P(G)$$

On cherche donc la probabilité qu'une gemme G se trouve à une position en sachant D les valeurs observées par le sonar. Cette probabilité est égale au la vraisemblance de ces positions observées en sachant les positions des gemmes et des probabilités précédentes.

$P(D | G)$ est donc calculé par la fonction de vraisemblance et $P(G)$ est calculé en multipliant toutes les probabilités précédentes des positions des gemmes (en les supposant indépendantes).

Finalement, α correspond au facteur de normalisation. Son but est de garantir que la somme des probabilités reste égale à 1. Dans le cours nous avons:

$$\alpha = \frac{1}{Z} = \frac{1}{P(\mathbf{e})} = \frac{1}{\sum_{\mathbf{q}} P(\mathbf{q}, \mathbf{e})}$$

Faisons correspondre cette formule à notre cas:

$$\alpha = \frac{1}{\sum_G P(G, D)}$$

avec

$$P(G, D) = P(D | G) P(G)$$

on a finalement

$$\alpha = \frac{1}{\sum_G P(D | G) P(G)}$$

3 Analyses de l'impact de paramètres sur la détection de gemmes

3.1 Impact du paramètre λ

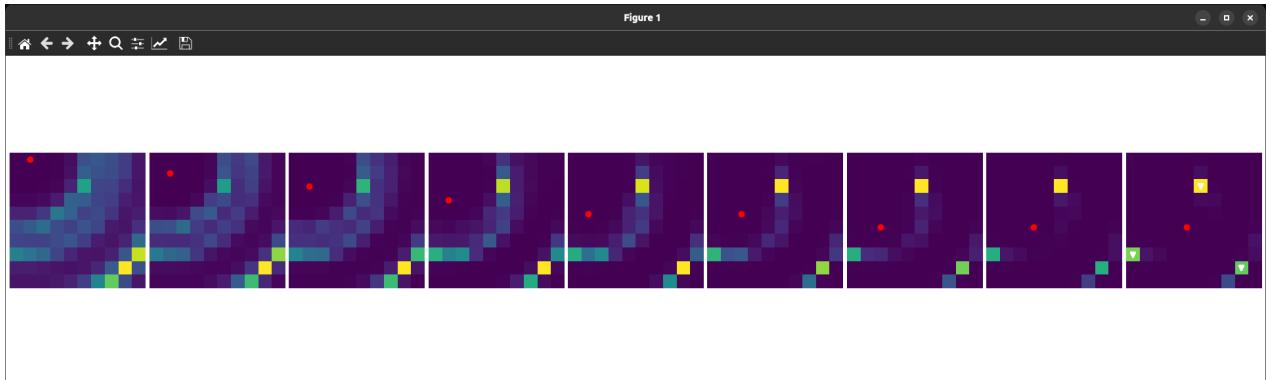


Figure 1: $\lambda = 1$

Comme expliqué plus haut, plus λ est petit, plus la vraisemblance est grande. En utilisant cette valeur plus élevée dans le théorème de Bayes, nous observons que la probabilité qu'une gemme G soit à une position donnée en sachant des positions observées augmente. En d'autres termes, notre recherche converge plus rapidement vers la solution.



Figure 2: $\lambda = 0.1$

Au contraire donc, si nous usons d'un λ plus élevé, la convergence vers la solution sera plus lente. Voici un exemple avec $\lambda = 10$; nous observons que le nuage d'observations n'est absolument pas dissipé à la fin de l'exécution du programme.

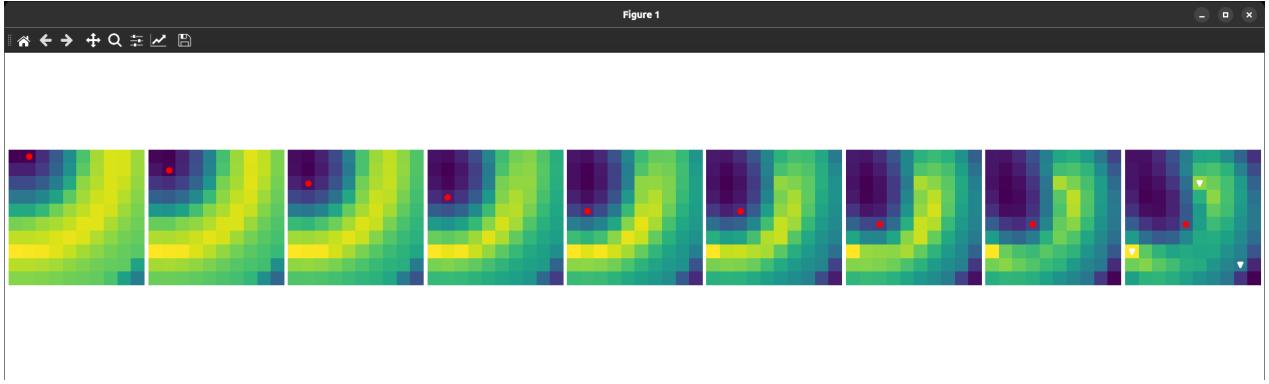


Figure 3: $\lambda = 10$

Attention cependant, cela ne veut pas dire que λ tendant vers 0 est un choix optimal. En effet, si les données sont bruitées, cela peut fortement impacter le résultat. Il faut donc trouver un λ optimisant la robustesse et la vitesse d’itération.

Pour mieux nous imager cela, nous pouvons faire une analogie avec les méthodes itératives numériques que nous avons apprises au cours de Calcul Formel et Numérique en BA2. Plus le pas d’itération est grand, plus nous convergeons vite vers une solution. Ceci est une source d’erreurs commune. En effet, le modèle itératif peut passer à côté de la solution sans jamais revenir (et donc donner un résultat totalement faux) ou encore revenir en arrière et osciller autour du résultat.

4 Amélioration de la méthode d’inférence

En théorie, sur base du cours 11, pour rendre la méthode d’inférence par énumération plus rapide, il faut faire usage du procédé d’inférence par élimination de variables pour éviter les facteurs répétés et ainsi réduire la charge de calcul.

Ceci est donc une optimisation certes, mais le problème reste NP-difficile.

5 Usage des LLMs

Comme précisé plus haut, nous avons fait usage de ChatGPT pour mieux comprendre le problème posé en addition au cours. Nous souhaitons préciser la nature de notre utilisation de cet outil. En effet, il est facile d’envoyer la question en *prompt*, récupérer la réponse et continuer sa vie.

Cette approche n’a aucun sens. En effet, l’étudiant procédant comme ceci n’apprend rien. C’est pourquoi, à chaque *prompt*, nous avons veillé à expliquer au mieux ce que nous comprenions. Ceci nous pousse à essayer d’exprimer le problème, auguisant ainsi notre compréhension. De plus, le LLM pourra fournir une réponse plus précise.

Même avec cette approche, il faut être vigilant. Nous sommes bien conscients que ChatGPT a tendance à toujours donner raison à son interlocuteur. C’est pourquoi il reste **primordial** de croiser ses sources, revenir sur le cours et consulter divers pages éducatives facilement trouvables sur Internet. Dans un monde idéal, un étudiant irait même en bibliothèque consulter un ouvrage de référence, mais nous ne sommes pas allés jusque là.

6 Conclusion

Ce second projet du cours d’Intelligence Artificielle nous a donc permis d’appliquer un nouvel aspect théorique de l’apprentissage machine vu en cours, en particulier les réseaux Bayésiens. Malgré quelques difficultés rencontrées en cours de route, nous avons fait de notre mieux pour nous tenir au procédé scientifique en expliquant au mieux notre démarche et arriver à un résultat satisfaisant.