

# **M1 Informatique – UE Projet**

## **Carnet de bord : les coulisses de la recherche documentaire**

---

---

PINAUD David - ANDROIDE

BIEGAS Emilie - ANDROIDE\_

**Sujet** : Algorithme Branch-And-Bound  
pour la résolution de diagramme  
d'influence

*Le mardi 23 février 2021*

## **Sommaire**

<b>Introduction</b>	<b>2</b>
<b>Les mots clés retenus</b>	<b>2</b>
<b>Descriptif de la recherche documentaire</b>	<b>3</b>
<b>Bibliographie produite dans le cadre du projet</b>	<b>4</b>
<b>Évaluation des sources</b>	<b>5</b>
Source [1]:	5
Source [5] :	6
Source [2] :	6

## I. Introduction

L'informatique décisionnelle est un champ d'étude permettant aux entreprises, aux hôpitaux, etc., de prendre des décisions optimales lorsqu'ils sont confrontés à des problèmes de décisions difficiles, avec beaucoup de critères par exemple. Cette discipline permet, par exemple et très particulièrement, d'aider à déterminer la maladie d'un patient présentant des symptômes propres à différents syndromes, et également d'aider à choisir sa médication.

Les diagrammes d'influences (ID) sont une représentation d'une situation de décision, utilisés pour modéliser un problème de décision et le résoudre par inférence.

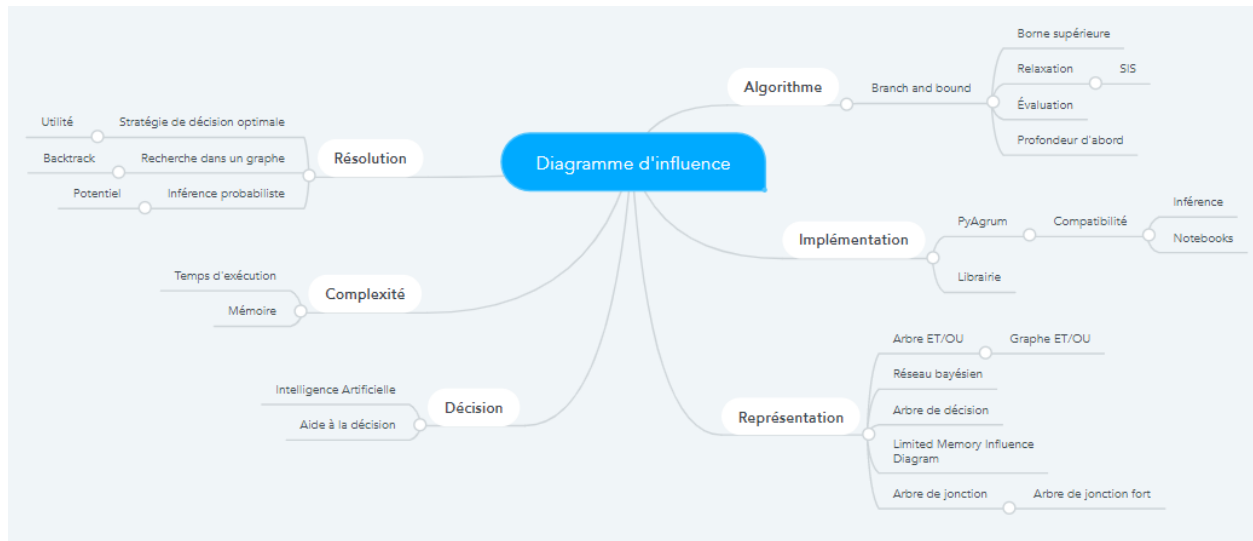
La plupart des algorithmes de résolution actuels convertissent l'ID vers un arbre de décision, une autre représentation d'un problème de décision, puis applique un algorithme pour le résoudre. Cette méthode a une complexité en temps et en espace explosive et ne permet que de résoudre des ID de petites tailles.

L'objectif de notre projet et de notre recherche est de trouver un algorithme permettant de résoudre de plus gros ID de façon efficace, basé sur une technique appelée le Branch-and-Bound (Séparation et Évaluation) puis de l'implémenter.

## II. Les mots clés retenus

- Diagramme d'influence/Influence Diagram
  - **Algorithme/Algorithm**
    - Branch and bound
    - Borne supérieure/Upper Bound
    - Relaxation/Relaxation
      - Ensemble suffisant d'information/Sufficient Information Set (SIS)
    - Évaluation/Evaluation
    - Profondeur d'abord/Depth-First
  - **Implémentation/Implementation**
    - PyAgrum
      - Compatibilité de l'inférence/Inference Compatibility
      - Compatibilité de notebooks/Notebook Compatibility
    - Librairie/Library
  - **Représentation/Representation**
    - Arbre ET-OU/AND-OR Tree
      - Graphe ET-OU/AND-OR Graph
    - Réseau bayésien/Bayesian Network
    - Arbre de décision/Decision Tree
    - Limited Memory Influence Diagram (LIMID)
    - Arbre de jonction/Junction Tree
      - Arbre de jonction fort/Strong Junction Tree
  - **Décision/Decision**

- Intelligence Artificielle/Artificial Intelligence
- Aide à la décision/Decision aid
- **Complexité/Complexity**
  - Temps d'exécution/Execution Time
  - Mémoire/Memory
- **Résolution/Resolution**
  - Stratégie de décision optimale/Optimal Strategy
    - Utilité/Utility
  - Recherche dans un graphe/Graph search
    - Retour en arrière/Backtrack
  - Inférence probabiliste/Probabilistic inference
    - Potentiel/Potential



### III. Descriptif de la recherche documentaire

Nos deux premiers articles scientifiques nous ont été fournis par notre encadrant. À partir de ces articles, nous avons exploré les citations qui nous ont paru pertinentes à notre objectif (recherche par rebonds). Nous avons également effectué des recherches par nom d'auteur lorsque cela était judicieux vis-à-vis du domaine de recherche de ce dernier et de nos besoins.

Quand une notion est discutée et que l'article ne possède pas de citations à ce sujet (et que les auteurs n'ont pas écrits de document sur cette notion), nous cherchons premièrement des articles à travers la barre de recherche dans l'onglet "bibliothèque de Sorbonne Université" sur le site de l'Université.

Si cela ne mène à aucun document pertinent, nous recherchons alors dans une des neuf bases de données concernant l'informatique proposées par le site de la bibliothèque, qui liste toutes les ressources en ligne.

Si cela ne mène à rien, nous recherchons les mots clés à travers Google Scholar en faisant bien attention à ce que les articles lus soient *peer-reviewed*.

Notre première approche à la recherche a été celle par rebond, elle est rapide, sûre et efficace, car elle aborde directement les raisonnements, théorèmes et concepts que l'on recherche lorsqu'on est en train d'étudier un article. Elle est toutefois assez limitée en contenu si l'auteur n'indique pas ou peu ses sources (d'ailleurs, même avec de bonnes citations, il est toujours pertinent et intéressant de chercher autre part, notamment étant donné que le nombre de documents cité est limité) et peut être inadapté à notre recherche puisque trop spécifique ou sortant du cadre de recherche. Avec cette approche et grâce à <http://www.researchgate.net>, nous avons pu retrouver les citations [1]–[5].

Notre seconde approche est donc la recherche par auteur, elle est également rapide, sûre et efficace à condition que le domaine de recherche de l'auteur en question soit en adéquation avec notre sujet. Elle est malgré tout assez limitée étant donné qu'un auteur écrit un nombre fini de documents qui sont, en général, sur divers sujets. Nous avons trouvé les articles [6] et [7] grâce à cette approche.

Ensuite, si on ne trouve rien de pertinent dans les citations du papier elles-mêmes, notre recherche se fait à travers la barre de recherche dans l'onglet "Bibliothèque". Celle-ci permet de chercher rapidement et de filtrer les résultats pertinents dans les bases de données reliées à la bibliothèque. Malheureusement, les articles sur notre sujet sont peu communs ou mènent à des liens cassés lorsqu'on veut lire en ligne sur des bibliothèques extérieures.

Viennent ensuite les bases de données dans les ressources en ligne. Ils sont nombreux et ne sont pas toujours pertinents. Nous avons exploré les bases de données dans la section "informatique" et n'avons trouvé aucun article pertinent à notre sujet, de même avec la base de données "europresse" qui donnait quelques articles de vulgarisation non souhaitables (en effet, notre sujet nécessite des ressources avec des explications précises et non vulgarisées).

Quand toutes les autres bases de données ou moteur de recherche ne permettaient pas de trouver des articles pertinents, on se tournait vers Google Scholar. C'est un outil efficace, car il centralise la plupart des publications, articles, revues, livres, etc. Il nous a en particulier aidé à trouver les articles [8]–[12]. Avec ce moteur de recherche (comme avec les autres, mais particulièrement celui-ci), il faut faire attention à ce que les articles soient bien *peer-reviewed* car il réfère à *tous* les documents, même ceux qui ne sont pas pertinents ou inexacts.

## IV. Bibliographie produite dans le cadre du projet

Norme IEEE (fait avec Zotero)

- [1] C. Yuan et E. Hansen, « Efficient Computation of Jointree Bounds for Systematic MAP Search. », *IJCAI Int. Jt. Conf. Artif. Intell.*, p. 1989, janv. 2009, [En ligne]. Disponible sur: [https://www.researchgate.net/profile/Changhe-Yuan/publication/220812142\\_Efficient\\_Computation\\_of\\_Jointree\\_Bounds\\_for\\_Systematic\\_MAP\\_Search/links/0fcfd51125a50c73e2000000/Efficient-Computation-of-Jointree-Bounds-for-Systematic-MAP-Search.pdf](https://www.researchgate.net/profile/Changhe-Yuan/publication/220812142_Efficient_Computation_of_Jointree_Bounds_for_Systematic_MAP_Search/links/0fcfd51125a50c73e2000000/Efficient-Computation-of-Jointree-Bounds-for-Systematic-MAP-Search.pdf).
- [2] D. Nilsson et S. Lauritzen, « Evaluating influence diagrams using LIMIDs », *Proc. Sixt. Conf. Uncertain. Artif. Intell.*, janv. 2013, [En ligne]. Disponible sur: [https://www.researchgate.net/profile/Steffen-Lauritzen/publication/234140127\\_Evaluating\\_influence\\_diagrams\\_using\\_LIMIDs/links/0deec51ec015546d69000000/Evaluating-influence-diagrams-using-LIMIDs.pdf](https://www.researchgate.net/profile/Steffen-Lauritzen/publication/234140127_Evaluating_influence_diagrams_using_LIMIDs/links/0deec51ec015546d69000000/Evaluating-influence-diagrams-using-LIMIDs.pdf).
- [3] A. Khaled, E. Hansen, et C. Yuan, « Solving Limited-Memory Influence Diagrams Using Branch-and-Bound Search », *Int. Symp. Artif. Intell. Math. ISAIM 2012*, sept. 2013, [En ligne]. Disponible sur: [https://www.researchgate.net/publication/257069324\\_Solving\\_Limited-Memory\\_Influence\\_D](https://www.researchgate.net/publication/257069324_Solving_Limited-Memory_Influence_D)

- iagrams\_Using\_Branch-and-Bound\_Search/link/545a5d2e0cf2c16efbbaae3d/download.
- [4] C. Yuan, X. Wu, et E. Hansen, *Solving Multistage Influence Diagrams using Branch-and-Bound Search*. 2012, p. 700.
  - [5] R. Dechter et R. Mateescu, « AND/OR search spaces for graphical models », *Artif. Intell.*, vol. 171, p. 73-106, févr. 2007, doi: 10.1016/j.artint.2006.11.003.
  - [6] S. Acid et L. M. de Campos, « An Algorithm for Finding Minimum d-Separating Sets in Belief Networks », in *Proceedings of the twelfth Conference of Uncertainty in Artificial Intelligence*, 1996, p. 22.
  - [7] D. D. Maua, C. P. de Campos, et M. Zaffalon, « Solving Limited Memory Influence Diagrams », *J. Artif. Intell. Res.*, vol. 44, p. 97-140, mai 2012, doi: 10.1613/jair.3625.
  - [8] D. R. Morrison, S. H. Jacobson, J. J. Sauppe, et E. C. Sewell, « Branch-and-bound algorithms: A survey of recent advances in searching, branching, and pruning », *Discrete Optim.*, vol. 19, p. 79-102, févr. 2016, doi: 10.1016/j.disopt.2016.01.005.
  - [9] R. A. Howard, J. E. Matheson, et Strategic Decisions Group, Éd., *Readings on the principles and applications of decision analysis*, Repr. Menlo Park, Calif: SDG, 1989.
  - [10] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, et S. D. Brown, « An introduction to decision tree modeling », *J. Chemom.*, vol. 18, n° 6, p. 275-285, 2004, doi: <https://doi.org/10.1002/cem.873>.
  - [11] D. Geiger, T. Verma, et J. Pearl, « d-Separation: From Theorems to Algorithms », in *Machine Intelligence and Pattern Recognition*, vol. 10, M. Henrion, R. D. Shachter, L. N. Kanal, et J. F. Lemmer, Éd. North-Holland, 1990, p. 139-148.
  - [12] David Kahle, Terrance Savitsky, Stephen Schnelle, et Volkan Cevher, « Junction tree algorithm », *Stat, Volume 631*, 2008.

## V. Évaluation des sources

### Source [1]:

Lors de notre lecture de l'article [3], nous avons besoin d'en savoir davantage sur les arbres de jonctions. Nous nous sommes donc orientés vers cet article cité traitant du sujet. Ce dernier a été publié en 2013 et est par conséquent assez récent. L'information trouvée dans ce document était très pertinente et répondait totalement à nos questions. Le niveau n'est ni trop général ni trop spécifique dans ce document, ce qui nous a permis d'améliorer notre compréhension du sujet.

Changhe Yuan et Eric A. Hansen sont les auteurs de cet article. Ils travaillent respectivement au département d'informatique à l'Université de New York et à l'Université du Texas à Dallas. Leurs domaines de recherche sont totalement en accord avec le thème de l'article. Ils sont, par conséquent, qualifiés pour parler du sujet et les informations présentes dans ce document nous paraissent donc légitimes. Le contenu de cet article est rigoureux, en effet, l'information est qualitative et objective, beaucoup de sources sont citées, et il ne semble pas y avoir d'erreur.

## Source [5] :

L'article [4] discutait de résolution d'un diagramme d'influence en utilisant un espace de recherche sous forme d'un arbre ET/OU. Il n'offrait pas de détails supplémentaires sur le sujet ni de citations pertinentes à la question. La source ([5]) en question est un état de l'art datant de 2005 qui proposait des définitions et des explications très claires sur cette représentation de l'espace de recherche. L'article a subi une mise à jour un an après sa publication en 2005, et a été accepté (*peer-reviewed*) peu après. Elle présente son contenu avec rigueur en définissant les objets et concepts utilisés tout en démontrant toutes les propositions et théorèmes proposés. L'article possède beaucoup de contenu (une trentaine de pages) spécialisé visant un public informé, et donc pertinent vis-à-vis de ce que l'on cherche.

Nous n'avons pas utilisé tout l'article, compte tenu de sa taille et de ce que l'on cherchait. Ce document a été produit dans l'objectif de faire un état de l'art sur les espaces de recherche en utilisant des graphes et arbres ET/OU sur des modèles graphiques probabilistes, des algorithmes utilisés pour faire de l'inférence (sur ces modèles spécifiquement) et de comparer leurs performances en temps et en espace. C'est donc un article de recherche sans biais, écrit par Rina Dechter, docteur en Informatique à l'Université de Californie ayant reçu des prix pour ses travaux (elle a d'ailleurs inventé les mots "apprentissage-profond"), et Robert Mateescu, aussi docteur en informatique à l'Université de Californie et post-doctorant à Caltech à l'époque de la publication de cet article.

## Source [2] :

Ce document nous a été fourni par notre encadrant qui est docteur en informatique et fait partie du laboratoire LIP6 de Sorbonne Université, il est donc très qualifié.

Cet article avait pour but de nous exposer le sujet du projet et constituait une bonne entrée en matière. Nous avons ainsi pu, suite à plusieurs lectures, comprendre les notions liées au projet, notamment grâce à un niveau qui n'est ni trop général ni trop spécifique dans ce document. Effectivement, cet article était à notre portée en termes de difficulté malgré la nécessité de relecture afin d'assimiler correctement toutes les notions.

Cette source est donc largement pertinente et constitue une partie intégrale de notre projet. Nous consultons d'ailleurs ce document régulièrement afin de se remettre les bases en tête ou d'approfondir quelques notions que nous aurions pu oublier.

Il a été écrit par Dennis Nilsson et Steffen L. Lauritzen, tous deux docteurs en mathématiques à l'Université de Aalborg au Danemark quand l'article a été écrit. Leur qualification est d'autant plus marquée qu'ils ont été primés pour leurs travaux. L'article n'a pas été révisé depuis sa publication en 2000. L'information est bien expliquée et démontrée prouvant la rigueur du contenu et ne démontre aucune erreur visible. C'est un article scientifique écrit pour un public informé et ne présentant aucun biais apparent.