

Discussion sur la causalité imparfaite et son application aux sciences expérimentales

Lucie Jacquin, Aurore Lomet, Jean-Philippe Poli

▶ **To cite this version:**

Lucie Jacquin, Aurore Lomet, Jean-Philippe Poli. Discussion sur la causalité imparfaite et son application aux sciences expérimentales. Rencontres francophones sur la Logique Floue et ses Applications, Université La Sorbonne, Oct 2021, PARIS, France. pp.187-190. cea-03406885

HAL Id: cea-03406885
https://cea.hal.science/cea-03406885

Submitted on 28 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Discussion sur la causalité imparfaite et son application aux sciences expérimentales

Lucie Jacquin Aurore Lomet Jean-Philippe Poli
Université Paris-Saclay, CEA, List

F-91120, Palaiseau, France, {prenom.nom}@cea.fr

Résumé :

La découverte de matériaux repose principalement sur l'analyse de données expérimentales obtenues par variation des paramètres des procédés de fabrication et en observant les propriétés des matériaux ainsi créés. L'Intelligence Artificielle (IA) a été appliquée à ce domaine, proposant des modèles pouvant prédire les performances des nouveaux matériaux. Nos travaux visent plutôt une application de l'IA en lien avec les experts des matériaux afin de leur fournir des connaissances extraites de leurs données. Une première étape d'importance est la détermination des liens causaux entre paramètres, propriétés et performances. Dans cet article, nous positionnons notre intérêt sur l'extraction de causalité dans le cadre de la logique floue.

Mots-clés :

Causalité, Sciences des matériaux, Causalité imparfaite, Logique floue.

Abstract:

Material discovery is mainly based on the analysis of experimental data obtained by varying the parameters of the manufacturing processes and by observing the properties of the created materials. Artificial Intelligence (AI) has been applied to this field, providing models that can predict the performance of new materials. Our work rather aims at an application of AI in connection with materials experts to provide them with knowledge extracted from their data. An important first step is to determine the causal relations between parameters, properties and performance. In this article, we position our interest on causality extraction within the framework of fuzzy logic.

Keywords:

Causality, Material science, Imperfect causality, Fuzzy logic.

1 Introduction

Cet article présente un positionnement de recherche à l'intersection de la recherche de liens causaux et de la logique floue. D'après [11], l'apprentissage automatique a besoin de modèles causaux de la réalité pour arriver au niveau de l'intelligence humaine. En effet, la compréhension des causalités est une caractéristique essentielle de notre compréhension du monde [1]. Par ailleurs, le cadre de la logique floue nous permet de manipuler des informations exprimées dans le langage naturel. Ainsi, une représentation floue de mécanismes causaux pourrait être très pertinente dans les applications d'extraction de connaissances nécessitant la compréhension humaine. Ce positionnement est motivé par une application de recherche de nouveaux matériaux pour laquelle l'interprétabilité des représentations par des humains est requise. Afin d'assister à la conception de matériaux, nous souhaitons être en mesure d'extraire des connaissances pertinentes et utiles pour les experts. Dans le cas des matériaux, les expérimentations ne sont pas toujours reproductibles, en particulier lorsqu'elles sont effectuées sur des machines de fabrication différentes. Il est donc nécessaire de prendre en compte une causalité imparfaite afin d'être en mesure malgré tout d'extraire des connaissances. En particulier, les connaissances qui sont généralement utilisées sont elles-mêmes vagues. Par exemple, "la viscosité d'un liquide est faible lorsque sa température est faible" (et inversement pour un gaz) ou encore "la lignine ne se gazéifie pas bien à basse température" [2].

2 Recherche de liens causaux

L'un des cadres les plus répandus pour décrire les mécanismes causaux sont les SCM (Structural Causal Models) [10]. Ils consistent le plus souvent en des équations structurelles et un graphe causal [8]. Les équations structurelles permettent de spécifier les effets causaux de chaque variable, tandis que le graphe cau-

sal est une interprétation causale d'un réseau bayésien $\mathbf{G} = (\mathcal{V}, \mathcal{E})$, où l'ensemble des nœuds $\mathcal{E}$ représente l'ensemble des variables en jeu dans le mécanisme causal et l'ensemble des arrêtes $\mathcal{V}$ représente l'ensemble des liens de cause à effet entre ces variables. Notons que le graphe causal est supposé vérifier la condition de Markov globale selon laquelle une variable est indépendante de ses non-effets conditionnellement à ses causes directes. Afin d'illustrer graphiquement cette condition, nous avons besoin de la notion de d-séparation. On dit que deux noeuds X et Y dans $\mathcal{E}$ sont d-séparés dans $\mathbf{G}$ par un sous-ensemble S de $\mathcal{E}$ si tous les chemins entre X et Y sont bloqués par S (voir [8] pour les détails sur la notion de chemins bloqués). La condition de Markov nous permet de lier un graphe causal à la distribution des données $\mathbf{D}$: si X et Y sont d-séparés par S dans $\mathbf{G}$ alors $X \perp\!\!\!\perp Y|S$ dans $\mathbf{D}$.

On distingue deux principales tâches dans l'étude des mécanismes causaux : celles qui visent à construire des relations causales à partir de données observées (découverte de causalité) et celles qui infèrent des effets causaux à partir d'un graphe causal donné (inférence causale). Puisque dans notre application, nous ne disposons pas de la liste exhaustive des liens de cause à effet présents, i.e., des arrêtes présentes, nous nous concentrons dans un premier temps sur une tâche de découverte de causalité. Le sujet de la découverte de causalité est très étayé dans la littérature [7, 8]. Il existe trois grandes familles d'approches pour la découverte de causalité, les méthodes basées sur des contraintes (CB), celles basées sur des scores (SB) et celles basées sur des *functional causal model* (FCM). Les méthodes appartenant à première famille exploitent les relations d'indépendance conditionnelle dans les données pour découvrir la structure causale sous-jacente. Ces méthodes font l'hypothèse appelée *Causal Faithfulness Assumption*, i.e., la réciproque de la condition de Markov : si $X \perp\!\!\!\perp Y|S$ dans D alors X et Y sont d-séparés par S dans $\mathbf{G}$. Grâce aux hypothèses de Markov et de *faithfulness*, nous avons une correspondance biunivoque entre les d-séparations dans le graphe et les indépendances conditionnelles dans la distribution. Dans la seconde famille d'approches SC, on cherche les graphes maximisant la qualité de l'ajustement à la distribution des données. Le score d'ajustement le plus utilisé à cette fin est le critère d'information bayésien (BIC) [14], mais d'autres scores ont été proposés [13]. Une fois le score défini, la recherche d'un graphe optimum pour la maximisation du score est effectuée par des méthodes heuristiques [3, 12]. Enfin la troisième famille de méthodes de découverte de causalité, FCM, vise à déterminer l'orientation des arrêtes dans son processus. Pour cela, ces méthodes utilisent l'hypothèse que les bruits des effets devraient être indépendants des causes. Par exemple, dans le cas où nous cherchons à identifier si X cause Y ou si Y cause X, le principe est d'envisager les deux possibilités, "$X = f(Y, \epsilon_X)$" et "$Y = f(X, \epsilon_Y)$", suivant des hypothèses sur la distribution des données et sur la relation fonctionnelle f, ϵ_X et ϵ_Y étant les bruits résultant pour ces deux configurations. Ces méthodes recherchent ensuite une asymétrie entre X et Y. Par exemple, si $Y \perp\!\!\!\perp \epsilon_X$ mais $X \not\perp\!\!\!\perp \epsilon_Y$, alors Y cause X.

3 Causalité imparfaite

Pour l'interprétation humaine, les liens de causalité sont souvent exprimés dans un langage vague. Cette imperfection peut concerner les définitions des causes et effets, par exemple, dans l'affirmation "*peu dormir* cause une fatigue inhabituelle". Ou bien, l'imperfection peut concerner le lien causal en lui même, comme dans l'affirmation, "dormir moins de 5 heures *peut provoquer* une fatigue inhabituelle". L'inconvénient des SCM est que si les noeuds ont un sens vague, la distribution ne peut être spécifiée de manière exacte [15]. Pour pallier ce problème et introduire l'imperfection dans la recherche de causalité, plusieurs méthodes alternatives ont été développées dont la plus courante est l'utilisation des *fuzzy cog-*

nitives maps de Kosko (KFCM) [9]. Les KFCM permettent de représenter et gérer la causalité imparfaite d'un système dynamique, i.e., présence de cycles dans la structure causale. Contrairement aux KFCM, nous nous positionnons dans un cadre acyclique puisque nous savons que les causes sont toujours des paramètres de fabrication et les effets, les propriétés des matériaux.

Une méthode alternative développée par [4, 6] consiste à formaliser les relations causales en se basant sur une couverture parcimonieuse des effets. Pour ce faire, les auteurs définissent des modèles de relations pour la causalité [5]. Cette méthode définit ainsi des relations possibles entre les causes et les effets sans en garantir la nécessité. Son avantage est de considérer des ensembles de causes possibles et de ce fait leurs interactions. Plus formellement, soit l'ensemble $\mathcal{X} = \{X_1, \ldots, X_k\}$ et $\mathcal{Y} = \{Y_1, \ldots, Y_k\}$, X_i peut causer Y_j mais pas nécessairement. Cependant, cette mise en évidence de liens causaux possibles ne semble pas primordiale dans le cadre de la fabrication de nouveaux matériaux puisque l'on souhaite garantir l'accessibilité des propriétés visées.

4 Positionnement et discussion

Notre objectif est d'utiliser les méthodes de recherche de liens causaux pour extraire des connaissances causales et les représenter de manière interprétable. Pour ce faire, nous proposons de travailler avec de nouvelles variables aléatoires décrivant l'appartenance à des ensembles flous définis au préalable. Puis, nous proposons de générer un graphe causal imparfait en appliquant des méthodes de recherche de causalité de l'état de l'art.

Plus formellement, considérons les variables $X_1, \ldots, X_p$, $p > 0$ en jeu dans la structure causale. Dans un premier temps, une phase d'extraction d'ensembles flous pertinents pour chaque variable doit être effectuée. Notons $A_i^1, \ldots, A_i^{m_i}$, $m_i \geq 0$, les ensembles flous obtenus à partir de la variable X_i. Ensuite, nous définissons une nouvelle variable aléatoire pour chaque ensemble flou par $Z_i^j = \mu_{A_i^j}(X_i)$, $j = 1, \ldots, m_i$ et $i = 1, \ldots, p$. Enfin, nous pouvons appliquer une méthode de découverte de causalité de type CB, SB ou FCM pour générer un graphe causal des variables $\{Z_i^j\}_{j=1,\ldots,m_i,\ i=1,\ldots,p}$ décrivant les valeurs prises par les degrés d'appartenance des ensembles flous. Considérons un exemple fictif où nous disposons d'observations pour les trois variables aléatoires suivantes : $X_1 :=$ "Paramètre de fabrication 1", $X_2 :=$ "Paramètre de fabrication 2", $X_3 :=$ "Propriété du matériau". La Figure 1 montre la définition de la nouvelle variable aléatoire Z_1^2 dans notre exemple fictif. Nous voyons qu'avec une observation x_1 de la variable X_1, nous pouvons obtenir une observation z_1^2 de la variable Z_1^2. Le graphe causal en Figure 2a est un exemple de résultat possible si on appliquait une méthode de recherche de causalité directement sur les variables disponibles. Dans ce graphe, nous avons les paramètres de fabrication X_1 et X_2 qui pointent tous les deux sur la propriété X_3. Cette représentation est juste mais ne nous informe pas sur les valeurs de paramètres et de propriétés réellement en jeu dans les relations causales. En revanche, avec un graphe causal, comme illustré en Figure 2b, généré à partir des nouvelles variables, nous serions en mesure d'extraire des informations plus précises sur la structure causale sous-jacente. Par exemple, le lien coloré nous indique que la valeur du degré d'appartenance du paramètre de fabrication 1 (X_1) à l'ensemble flou "proche de 0" a un effet direct sur la valeur du degré d'appartenance de la propriété (X_3) à l'ensemble flou "bon". Autrement dit, la proximité de "X_1 à la valeur 0" a un effet direct sur la qualification de X_3 dans l'état "bon". La prise en compte d'effets joints (entre plusieurs causes) pourrait être également envisagée avec des variables aléatoires considérant des ensembles flous d'éléments joints. Néanmoins, cette considération ne semble réalisable et interprétable que pour un faible nombre de facteurs initiaux.

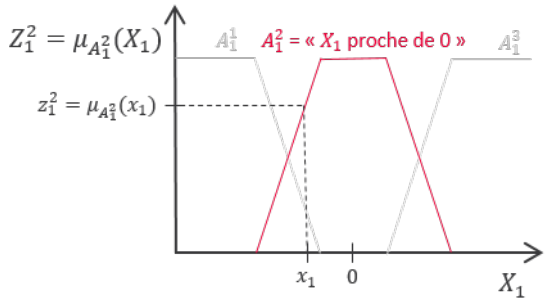
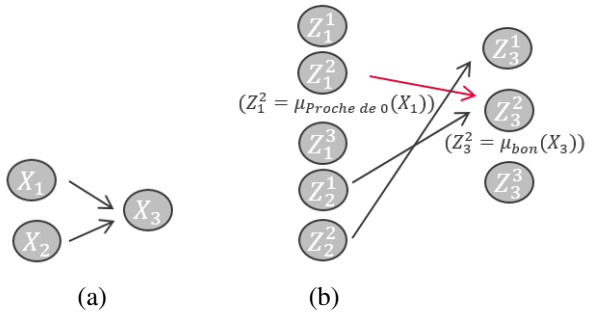


Figure 1 – Définition de la nouvelle variable Z_1^2



(a) (b)

Figure 2 – Exemples fictifs de graphes causaux construits, avec les variables d'origine (a), avec les nouvelles variables (b)

La première perspective de cette discussion nous semble être la mise en application de la procédure envisagée sur des données réelles afin de retrouver automatiquement les liens causaux entre les paramètres de fabrication et les propriétés du matériau. Les résultats pourront être présentés devant des experts afin d'être validés.

Une fois ces liens causaux trouvés, nous nous intéresserons à l'induction de règles. Cela nécessitera d'identifier les différents types de liens, comme par exemple la gradualité, qui détermineront la forme de la règle. Cela peut nécessiter également de découvrir automatiquement les modalités associées aux différentes variables.

Enfin, nous serons en mesure d'évaluer la méthode globale sur des données expérimentales issues de différents cas d'usage comme les batteries électriques et les cellules photovoltaïques.

Remerciements :

Ces travaux sont financés par le programme transverse de compétences "Matériaux et Procédés" du CEA.

Références

[1] Cristina Puente Agueda. Causality in sciencie. *Pensamiento Matemático*, (1) :12, 2011.

[2] Prabir Basu. Chapter 8 - design of biomass gasifiers. In Prabir Basu, editor, *Biomass Gasification, Pyrolysis and Torrefaction (Third Edition)*, pages 263–329. Academic Press, third edition edition, 2018.

[3] David Maxwell Chickering. Optimal structure identification with greedy search. *Journal of machine learning research*, 3(Nov) :507–554, 2002.

[4] Didier Dubois and Henri Prade. Fuzzy relation equations and causal reasoning. *Fuzzy sets and systems*, 75(2) :119–134, 1995.

[5] Didier Dubois and Henri Prade. An overview of ordinal and numerical approaches to causal diagnostic problem solving. *Abductive reasoning and learning*, pages 231–280, 2000.

[6] Didier Dubois and Henri Prade. A glance at causality theories for artificial intelligence. In *A Guided Tour of Artificial Intelligence Research*, pages 275–305. Springer, 2020.

[7] Clark Glymour, Kun Zhang, and Peter Spirtes. Review of causal discovery methods based on graphical models. *Frontiers in genetics*, 10 :524, 2019.

[8] Ruocheng Guo, Lu Cheng, Jundong Li, P Richard Hahn, and Huan Liu. A survey of learning causality with data : Problems and methods. *ACM Computing Surveys (CSUR)*, 53(4) :1–37, 2020.

[9] Bart Kosko. Fuzzy cognitive maps. *International journal of man-machine studies*, 24(1) :65–75, 1986.

[10] Judea Pearl. *Causality*. Cambridge university press, 2009.

[11] Judea Pearl. Theoretical impediments to machine learning with seven sparks from the causal revolution. *arXiv preprint arXiv :1801.04016*, 2018.

[12] Joseph Ramsey, Madelyn Glymour, Ruben Sanchez-Romero, and Clark Glymour. A million variables and more : the fast greedy equivalence search algorithm for learning high-dimensional graphical causal models, with an application to functional magnetic resonance images. *International journal of data science and analytics*, 3(2) :121–129, 2017.

[13] Teemu Roos, Tomi Silander, Petri Kontkanen, and Petri Myllymaki. Bayesian network structure learning using factorized nml universal models. In *2008 Information Theory and Applications Workshop*, pages 272–276. IEEE, 2008.

[14] Gideon Schwarz et al. Estimating the dimension of a model. *Annals of statistics*, 6(2) :461–464, 1978.

[15] Alejandro Sobrino. Imperfect causality : Combining experimentation and theory. In *Combining Experimentation and Theory*, pages 371–389. Springer, 2012.