

Thèse : Causalité et probabilités : réseaux bayésiens, propensionnisme

Résumé synthétique :

Les théories probabilistes de la causalité sont apparues au milieu du vingtième siècle, essentiellement pour prendre en compte ceci que la causalité n'est pas une relation de nécessité. Elles se sont développées depuis lors et ont acquis une place centrale dans le champ des théories contemporaines de la causalité. Ma thèse traite de questions qui se posent en l'état actuel du développement des théories probabilistes de la causalité. Il apparaît très rapidement que ces questions sont très différentes selon qu'on s'intéresse à la causalité générique, entre propriétés, ou à la causalité singulière, entre événements qui adviennent effectivement. Pour la première, on peut considérer que l'analyse conceptuelle de ses rapports avec les probabilités est achevée, laissant ouvertes des questions d'ordre épistémologique. Plus exactement, la première partie de mon travail vise à évaluer la contribution de méthodes graphiques récentes à l'inférence aux causes génériques à partir de données statistiques. De façon sensiblement différente, la question du rapport entre la causalité singulière et les probabilités n'est pas complètement réglée du point de vue conceptuel. La seconde partie de mon travail aborde cette question à partir d'une analyse de la relation entre la causalité et la théorie propensionniste des probabilités.

Résumé étendu :

L'analyse de la causalité a pris un tournant au milieu du vingtième siècle, quand il est devenu évident que la causalité n'est pas une relation de nécessité. Pour le dire autrement, il est apparu que de nombreuses causes ne sont pas des conditions suffisantes pour leurs effets. Ainsi, fumer cause le cancer alors même que certains fumeurs ne développent jamais de cancer. De même, la sélection naturelle peut être considérée comme une cause de l'augmentation de la fréquence relative d'un trait dans une population alors même qu'elle ne rend pas cette augmentation de fréquence nécessaire. De façon sensiblement différente, le cambriolage des locaux du Parti Démocrate a causé la démission du président Nixon, mais il aurait pu ne pas avoir cet effet. Enfin, il est possible que l'événement décrit par la proposition « Pierre a placé le compteur Geiger à proximité d'un échantillon radioactif le 3 décembre 2008 à 8h12 » ait causé le déclenchement du compteur Geiger le même 3 décembre à 8h18, mais en aucun cas il ne l'a pas rendu nécessaire.

Les théories probabilistes de la causalité qui apparaissent dans les années 1960 visent précisément à prendre en compte ceci que la causalité n'est pas une relation de nécessité. En effet, l'idée qu'on trouve au fondement de ces théories est la suivante : une cause, si elle ne rend pas ses effets nécessaires, les rend plus probables. Un peu plus formellement, l'idée est de caractériser la relation de cause à effet entre A et B au moyen de l'inégalité suivante : $p(B | A) > p(B)$. Les théories probabilistes de la causalité se sont développées à partir de cette idée séminale. Mon travail de thèse examine des questions qui se posent en l'état actuel du développement de ces théories.

Il apparaît dès l'introduction de la thèse que ces questions sont très différentes selon qu'on s'intéresse à la causalité entre des propriétés (par exemple celle de fumer et celle de développer un cancer du poumon) ou à la causalité entre des événements singuliers (par exemple l'événement décrit par la proposition « Pierre a placé le compteur Geiger à proximité d'un échantillon radioactif le 3 décembre à 8h12 » et le déclenchement du compteur Geiger). Dans le premier cas, celui de la causalité générique, on peut considérer achevée l'analyse du rapport entre les concepts de causalité et de probabilité. Les questions qui se posent sont donc des questions de nature épistémologique,

relatives plus précisément à la mise au jour de relations de cause à effet entre propriétés. Dans le second cas, celui de la causalité singulière, je soutiens qu'il reste à proposer une théorie probabiliste de ce que j'appelle « causalité actuelle », c'est-à-dire de la causalité considérée comme une relation physique entre le contenu de certaines zones d'espace-temps. Le plan de la thèse découle de l'affirmation selon laquelle la causalité générique et la causalité singulière ne sont pas dans des situations analogues relativement aux théories probabilités de la causalité.

La première partie de ma thèse traite de questions d'épistémologie de la causalité générique. Plus précisément, l'analyse porte sur l'inférence aux causes génériques en tant qu'elle prend appui sur les réseaux bayésiens. Les réseaux bayésiens sont des graphes probabilistes qui ont été introduits en intelligence artificielle dans les années 1980. Plus récemment, ils ont permis de définir de puissants algorithmes d'inférence causale que les scientifiques utilisent de plus en plus volontiers. Je considère que ces algorithmes sont les meilleurs candidats au titre d'outils d'inférence issus de la caractérisation de la causalité par les théories probabilistes et je me propose d'évaluer leur contribution à l'inférence aux causes génériques.

Le chapitre 1 est consacré à présenter les réseaux bayésiens en général, et plus particulièrement les réseaux bayésiens causaux (ou, plus exactement, causalement interprétés). Je discute les hypothèses que la notion même de réseau bayésien causal véhicule relativement à la causalité et à son rapport avec les probabilités. Je fais état de l'existence de violations de ces hypothèses, puis j'explique comment on peut les rétablir au moyen d'artifices dans la représentation des systèmes qu'on considère. La conclusion principale est la suivante : dans le contexte d'inférence causale, le statut de ces hypothèses n'est pas tel que les artifices qui permettent de les rétablir peuvent être effectivement utilisés.

Le chapitre 2 porte sur le critère de causalité qui est mis en œuvre quand on utilise les méthodes d'inférence aux causes fondées sur les réseaux bayésiens. Plus précisément, une fois ce critère mis au jour, il s'agit de déterminer comment il se situe par rapport aux théories probabilistes de la causalité générique. Dans un premier moment, je mets en évidence deux différences entre ce que vise ce critère d'une part et l'objet des théories probabilistes de la causalité d'autre part. Dans un second moment, je m'efforce de mener à bien la comparaison non plus seulement des objets mais des analyses elles-mêmes. Pour cela, je construis un objet qui peut être analysé à la fois au moyen du critère de causalité véhiculé par les réseaux bayésiens causaux et au moyen des théories probabilistes de la causalité. Relativement à cet objet, j'établis que le critère de causalité véhiculé par les réseaux bayésiens causaux prend bien place dans le champ des théories probabilistes de la causalité, et je l'y situe précisément. Conjointe à la différence entre les objets des deux types d'analyses, cette situation explique que les réseaux bayésiens permettent d'inférer des causes génériques là où nos meilleures théories probabilistes de la causalité ne peuvent pas être utilisées pour les reconnaître.

Le chapitre 3 est un chapitre de méthodologie de l'inférence aux causes dans les sciences sociales au sens large – c'est-à-dire dans le domaine pour lequel les réseaux bayésiens sont un outil d'analyse approprié. Plus exactement, il s'agit de déterminer ce que les algorithmes fondés sur les réseaux bayésiens changent à l'inférence causale, et en particulier si l'existence de violations des hypothèses corrélatives de la notion de réseau bayésien causal implique simplement qu'il faut réserver l'inférence aux causes fondées sur les réseaux bayésiens aux systèmes pour lesquels ces hypothèses ne sont pas violées. L'approche adoptée est comparative : je m'efforce de déterminer en quoi les méthodes d'inférence aux causes fondées sur les réseaux bayésiens se distinguent des méthodes hypothético-déductives traditionnellement utilisées en sciences sociales. La principale conclusion est la suivante : ce qui est spécifique de l'inférence aux causes fondée sur les réseaux bayésiens est le caractère déductif de l'inférence quand elle est justifiée, et corrélativement le statut des difficultés soulevées par la possibilité que les hypothèses corrélatives des réseaux bayésiens causaux soient violées. Plus explicitement, dans le cadre de l'inférence aux causes fondée sur les réseaux bayésiens et seulement dans ce cadre, la possibilité de ces violations implique que la

conclusion de l'inférence causale est toujours suspecte d'être fautive (et ce alors même qu'elle peut fort bien être vraie). Ce résultat acquis, j'explique comment les algorithmes qui recourent aux réseaux bayésiens peuvent néanmoins être mobilisés dans un contexte d'inférence causale si on les intègre aux procédures hypothético-déductives.

Le chapitre 4 traite une question orthogonale à celles qui sont abordées dans les chapitres 2 et 3. En effet, il ne s'agit plus de caractériser l'inférence causale fondée sur les réseaux bayésiens, mais de délimiter le domaine au sein duquel elle est valide, c'est-à-dire de délimiter le domaine au sein duquel le critère de causalité véhiculé par les réseaux bayésiens est adéquat. Plus précisément, le chapitre cherche à évaluer un argument en faveur de la thèse selon laquelle le caractère déterministe ou indéterministe d'un système ne change rien relativement à la satisfaction de la condition de Markov causale – une composante essentielle du critère de reconnaissance des causes que les réseaux bayésiens véhiculent. L'argument discuté est rejeté à partir d'une analyse de la distinction entre variables exogènes et variables d'erreur. Positivement, je définis un sens précis auquel le déterminisme est plus favorable que l'indéterminisme pour la satisfaction de cette condition.

La seconde partie de la thèse contribue à l'analyse du rapport entre le concept de probabilité et celui de causalité actuelle. Cette question est abordée à partir de la philosophie des probabilités. En effet, une seule théorie des probabilités est susceptible de fonder une caractérisation de la causalité actuelle définie comme relation *physique* entre des événements *singuliers*. Il s'agit de la théorie propensionniste formulée par Popper au milieu des années 1950.

Le chapitre 5 consiste en une présentation approfondie de la théorie propensionniste des probabilités. Selon cette théorie, les probabilités mesurent des entités qui tendent à faire advenir les événements possibles et que Popper appelle « propensions ». J'examine les corrélats ontologiques, épistémologiques et métaphysiques de cette proposition, qui se révèlent non triviaux. Tout au long du chapitre, l'accent est mis sur l'analyse du rapport entre les propensions et les causes, et plus précisément sur la question de savoir si et en quel sens les propensions sont des causes.

Les analyses du chapitre 5 portent sur les probabilités absolues, de la forme $p(B)$, de la même manière que la théorie propensionniste des probabilités telle qu'elle est développée par les popperiens. Or, l'idée qu'on trouve au fondement des théories probabilistes de la causalité engage également la notion de probabilité conditionnelle, de la forme $p(B | A)$. Dans ces conditions, le chapitre 6 consiste dans un premier temps à étendre la théorie propensionniste aux probabilités conditionnelles. Proposer une telle extension revient à proposer une solution au « paradoxe de Humphreys », selon lequel il ne saurait y avoir d'interprétation propensionniste des probabilités conditionnelles. La seconde partie du chapitre vise à déterminer quel est le rapport entre les probabilités conditionnelles telles que je propose de les interpréter et la causalité actuelle. Les résultats obtenus sont au nombre de deux. En premier lieu, dans le cadre de la théorie proposée pour les probabilités conditionnelles, la notion d'augmentation de probabilités mobilisée par les théories probabilistes pour caractériser la causalité est comparable (et comparée) à la causalité singulière indéterministe telle qu'elle est définie par Lewis. En second lieu, je montre qu'il convient de distinguer la notion de dépendance causale que la théorie des probabilités conditionnelles proposée permet de définir de la notion de production causale qu'on trouve au cœur du propensionnisme popperien – et qui a été discutée dans le chapitre 5.