

## Réflexions sur la causalité

J'ai lu récemment *L'esprit dans un monde physique*, un recueil de quatre conférences par Jaegwon Kim sur le problème métaphysique que pose l'action de l'esprit sur le monde réel, le tout dans un cadre matérialiste et avec une absence totale de mysticisme que j'apprécie particulièrement.

Le problème principal est le suivant : si nous disposons de lois fondamentales qui décrivent complètement l'univers depuis le niveau subatomique, alors ces lois décrivent et prédisent tous les événements et phénomènes, y compris les phénomènes mentaux, comme la pensée « je veux me gratter la tête ». Dans ces conditions, comment peut-on dire que cette pensée a été la cause du mouvement ? La cause du mouvement est l'ensemble d'influx nerveux dans le cerveau. C'est ce que Jaegwon Kim et beaucoup d'autres appellent le *problème de la causalité mentale*.

Jaegwon Kim cite un argument important qui n'est pas le sien et qui façonne tout son discours. L'étude des phénomènes mentaux, des pensées, n'est pas fondamentalement différente de l'étude d'autres phénomènes qui apparaissent à l'échelle macroscopique comme conséquences des lois fondamentales de l'univers. Or s'il est évident que la causalité existe au niveau de la chimie ou de la mécanique, par exemple, elle doit forcément exister au niveau mental. Dès lors, le problème n'est pas de décider si cette causalité mentale existe mais d'expliquer comment elle peut exister.

Mais ma réaction face à une telle évidence est l'opposé : si quelque chose qui semble évident cesse de fonctionner quand on le pousse dans ses retranchements, j'ai tendance à penser qu'il y a en fait quelque chose de subtil caché derrière l'évidence.

Et ici, il ne faut pas réfléchir très longtemps pour trouver des situations communes où la notion de causalité se montre bien plus subtile qu'il n'y paraît. Dans des débats de société, par exemple, il est fréquemment nécessaire de rappeler à son contradicteur que le fait que deux événements qui arrivent systématiquement ensemble ou en succession ne veut pas forcément dire qu'il y a causation entre les deux, ce qu'on dira parfois en latin, *post hoc ergo propter hoc* ou *cum hoc ergo propter hoc*. Et pourtant, on est bien embarrassé si on cherche à produire des critères objectifs pour évaluer quand cette déduction est valide.

Pour fixer les idées, je prendrai un exemple en histoire. Un exemple très simplifié, parce que le but est d'expliquer mes réflexions sur une idée qui n'a au fond rien à voir avec l'histoire. Certains disent que la première guerre mondiale a été causée par l'assassinat de l'archiduc. D'autres insistent que la vraie cause est le nationalisme généralisé des sociétés occidentales à l'époque. Comment départager ces deux affirmations ? Ou même simplement : comment leur donner un sens ?

Jaegwon Kim évoque quelques tentatives de ses pairs pour clarifier un peu



la question. En particulier, il évoque une formulation « contrefactuelle » : si l'archiduc n'avait pas été assassiné, la première guerre mondiale n'aurait pas eu lieu. Avec le vocabulaire de la logique mathématique, cette formulation est une contraposée, plus précisément la contraposée de la réciproque de l'implication évidente, l'implication entre la cause et la conséquence.

Cependant, on aimerait également avoir l'implication directe : si la cause se produit, alors la conséquence se produit, c'est la moindre des choses. Puisque nous avons également la réciproque : si la cause ne se produit pas, alors la conséquence non plus, c'est que nous avons une équivalence.

Ce raisonnement montre que cette caractérisation de la causalité est incomplète : le connecteur d'équivalence est symétrique, la relation de causation ne l'est pas, on peut même arguer que le terme antisymétrique serait adapté. Il manque un critère pour distinguer le sens de la causalité, les causalités indirectes, les simples coïncidences.

Mais cette caractérisation a une autre lacune, plus subtile mais plus fondamentale je pense. On néglige l'implication directe ici parce qu'on ne part pas d'une situation hypothétique, on part de la situation réelle : l'archiduc a été assassiné, la première guerre mondiale a eu lieu, il n'est pas nécessaire de mettre un *si* devant la phrase.

Le problème mis en évidence ici est qu'on a négligé de préciser les quantificateurs qui s'appliquent aux variables du problème. Pour toute histoire hypothétique, le fait que cette histoire comporte l'assassinat de l'archiduc est équivalent au fait que cette histoire comporte une première guerre mondiale. Certes, mais quelles sont les histoires alternatives acceptables ?

Puisque j'ai été amené à parler d'implication, d'équivalence et de quantification, il semble naturel de s'intéresser au domaine d'où viennent ces notions : les mathématiques. Et de fait, on l'a déjà fait avant moi : l'article Une méditation sur le nombre 24 et la causalité en mathématiques de David Madore a été pour moi une source de réflexions sur la notion de causalité bien avant le livre de Jaegwon Kim.

Pour résumer à ma façon : on peut être tenté de se demander si c'est à cause du théorème de Pythagore que le triangle 3-4-5 est rectangle ou si c'est le contraire, mais il est difficile de donner un sens à cette affirmation : le théorème de Pythagore est vrai, le triangle 3-4-5 est rectangle, c'est tout. Pourtant, certaines coïncidences remarquables laissent soupçonner des explications plus profondes que des démonstrations calculatoires. L'article le dit mieux que moi.

Dans l'exemple de la première guerre mondiale, on peut imaginer une histoire alternative assez facilement. Au contraire, il n'y a pas de mathématiques alternatives, il n'y a que ce qu'on observe à travers des raisonnements parfaitement rigoureux.



Cet argument me donne envie de dire que l'idée des « contrefactuels » n'est qu'une approximation de l'idée intuitive que nous avons de la causalité. Je vais l'examiner un peu plus en détails.

Les partisans de la cause nationaliste réfutent la cause assassinat avec l'argument que, si l'archiduc n'avait pas été assassiné, autre chose aurait déclenché la première guerre mondiale. Donc pour défendre cette affirmation, il faut se limiter à des histoires alternatives très proches de l'histoire réelle, celles où aucun autre personnage important n'a été assassiné. Enfin, très proches, seulement jusqu'au moment où la première guerre mondiale ne se produit pas, parce que c'est un assez gros événement. Bref, ce sont des hypothèses assez forcées.

D'ailleurs, comment supposer que l'assassinat n'a pas eu lieu ? Gavrilo Princip avait des raisons pour son acte. La cause de la première guerre mondiale avait elle-même des causes, qui avaient elles-mêmes des causes, etc., jusqu'à la nuit des temps.

Je vais donner un dernier argument, extrême : dans l'histoire alternative racontée par Tolkien, la première guerre mondiale s'appelle la guerre de la grande colère, mais aucun archiduc n'a été assassiné. Cet argument est ridicule ? Bien sûr, c'est son objet.

Tout ça pour dire que quand on essaie de modéliser l'idée de causalité en histoire par des hypothèses sur des histoires alternatives, on le fait en supposant que ces histoires alternatives son assez similaires à l'histoire réelle sur les points importants pour le raisonnement, mais assez dissimilaires pour permettre de conclure, et ces hypothèses s'avèrent au final très artificielles.

Le problème dans cette démarche, c'est que nous n'avons aucune idée de la topologie et de la géométrie de l'espace des histoires possibles, ou mieux, des histoires plausibles.

Faute de connaître cette géométrie en général, nous devons nous fier à notre intuition pour essayer de déterminer, quand on affirme au sujet d'une situation historique, que les mêmes causes amènent les mêmes circonstances, ce qui exactement est recouvert par ce terme vague, « même ». Ce n'est pas en soi un problème, notre intuition a un certain degré de fiabilité en ce qui concerne les situations réelles, qui nous sont familières, puisqu'elle a justement été optimisée pour les traiter. Nous pouvons nous appuyer dessus pour prendre des décisions politiques, par exemple, surtout si on la complète par d'autres arguments et critères.

Mais une compréhension basée sur l'intuition n'est pas généralisable. Elle ne peut pas servir de base à une discussion philosophique.

Je pense personnellement que l'histoire est chaotique, au sens mathématique du terme : des changements mineurs dans les circonstances peuvent rapidement conduire à des résultats radicalement différents. Pas forcément n'importe quel résultat, ça n'exclue pas de dégager des tendances globales, mais on ne pourra pas



prévoir comment ces tendances seront implémentées. On peut penser à l'attracteur de Lorenz : on ne peut pas prédire où sera le point à un instant donné, mais on sait qu'il sera quelque part sur l'attracteur.

Je n'ai pas d'argument fort pour étayer cette affirmation, à part invoquer la sensibilité aux conditions initiales de phénomènes qui influencent l'histoire sans eux mêmes être historiques : la gamétogenèse, la météo, etc. Mais je n'ai rien pour étayer un chaos à court terme provoqué par des mécanismes spécifiquement historiques. Cependant, la sagesse populaire, avec la comptine « faute d'un clou », semble indiquer que j'ai peut-être raison.

Si ma conjecture est juste, alors on peut dire que la causalité en histoire n'existe pas vraiment, puisqu'on ne peut jamais retrouver exactement les mêmes circonstances de manière réaliste.

Cette situation me rappelle un peu la notion d'espèce en biologie. On considère en général que deux individus sont de la même espèce s'ils sont interfertiles. Pour que cette définition conduise bien à une partition du vivant en espèces, il faut que la relation d'interfertilité soit une relation d'équivalence, en particulier transitive. On fera une pirouette pour négliger que certains individus ne sont pas fertiles, et aussi que souvent un individu ne peut être interfertile qu'avec la moitié de son espèce ; ces détails ne sont pas graves. Ce qui l'est, c'est que la relation n'est pas toujours transitive globalement : on peut trouver des chaînes de sous-espèces qui sont interfertiles entre paires voisines mais pas d'un bout à l'autre.

Ça ne veut pas dire que la notion d'espèce n'est pas intéressante. Elle permet de faire énormément d'affirmations biologiques utiles qui seront effectivement vraies si on les vérifie expérimentalement. Mais il faut se rappeler que ce n'est qu'une approximation pratique. Il est fort possible que la même affirmation soit vraie pour la notion de causalité.

J'ai un dernier point à développer avant de conclure. Les lois de la physique de bas niveau sont souvent exprimées sous la forme d'équations différentielles ordinaires : des formules qui disent comment un système évolue en fonction de son état. Intuitivement, on peut penser que c'est une manifestation de la causalité : l'état actuel du système est la cause de son état futur.

Mais il faut se rappeler que ces lois de bas niveau sont, en l'état actuel des connaissances, toujours réversibles : on peut aussi déduire l'état antérieur du système de son état actuel. Donc l'état actuel serait la cause de l'état passé, ce qui est absurde. Et pourtant, ces lois peuvent donner naissance, par le biais des statistiques, à des lois irréversibles, en particulier la fameuse thermodynamique.

Non seulement les lois de bas niveau sont réversibles, mais en plus elles peuvent être formulées sous une forme variationnelle, le principe de moindre action, qui énonce une condition sur l'ensemble de l'histoire du système, prise dans sa globalité. Dans cette formulation, il n'y a plus vraiment de causalité.



Et c'est sur ça que je vais m'appuyer pour conclure : notre Univers est une solution particulière d'une équation différentielle globale extrêmement compliquée. Nous voyons des phénomènes se produire dans cette solution, nous soupçonnons que ces phénomènes obéissent à une logique, à des lois applicables aux autres solutions similaires, mais nous n'avons pas encore pu énoncer ces lois rigoureusement. Pour nous en approcher, nous nous appuyons sur des critères intuitifs, et la causalité est un de ces critères, mais rien ne permet d'affirmer qu'il est effectivement pertinent ni qu'il pourra se traduire directement en une loi rigoureuse.

Dans Les planètes, les aimants et la conscience, j'é mets la conjecture un peu folle que l'explication de la réduction du paquet d'onde est à chercher au niveau subjectif : la conscience ne pourrait se manifester, observer, que des mondes où les paquets d'onde sont réduits, des mondes presque exclusivement classiques. Je peux prolonger cette conjecture pour la notion de causalité : peut-être la conscience telle que nous la connaissons, une forme de conscience ancrée dans la temporalité, ne peut-elle se manifester, ne peut-elle observer que des mondes causaux.

*Publié le 13 juin 2019*

