de celle du temps interdit une construction inverse, consistant à fonder la direction du temps sur celle de la causalité – idée que certains philosophes trouvent séduisante. Et au plan méthodologique, la proposition de P. Suppes n'est guère éclairante : autant les informations probabilistes sont souvent plus faciles à obtenir que les informations causales, autant ce n'est pas toujours le cas des informations temporelles. Par exemple, il semble clair que le prénom est antérieur à la mention au baccalauréat ; mais quand on cherche à relier un taux de chômage élevé avec une croissance économique faible, laquelle des deux propriétés est antérieure à l'autre ?

Dans une seconde famille de théories probabilistes de la causalité, on cherche plutôt à préciser l'idée selon laquelle les causes rendent leurs effets plus probables. Être une cause de E, c'est rendre E plus probable dans les différentes sous-populations définies par les autres variables pertinentes. Autrement dit, le prénom Irène est la cause de l'obtention d'une mention « Très bien » si et seulement si « Irène » rend cette mention plus probable une fois que toutes les autres informations pertinentes ont été prises en compte. Supposons pour simplifier que le capital culturel des parents soit la seule information de ce type (ce n'est en réalité pas le cas) ; cela signifie que « Irène » serait la cause de la mention « Très bien » si et seulement si ce prénom rendait plus probable cette mention à la fois dans la sous-population des candidates culturellement favorisées et dans celle des autres candidates.

Est-il possible de caractériser de façon générale de telles sous-populations ? Le capital culturel a un effet causal sur la mention au baccalauréat, et le prendre en compte suffit ici à garantir l'analyse en termes de corrélation positive parce que nous avons supposé que, le prénom mis à part, il constitue la seule information pertinente. Autrement dit, nous avons supposé que, le prénom mis à part, le capital culturel est la seule autre cause possible de la mention. Et nous avons distingué seulement deux sous-populations homogènes vis-à-vis du capital culturel, celle des candidates culturellement favorisées et celle des candidates qui ne le sont pas. L'analyse pourrait être améliorée en décomposant plus finement la population en fonction du capital culturel : « Irène » causerait la mention « Très bien » si et seulement si ce prénom rendait la mention « Très bien » plus probable dans toutes les sous-populations définies par les différentes valeurs du capital culturel.

L'IDÉE INTRODUITE EN 1979
par Nancy Cartwright a conduit à une analyse de la causalité que les philosophes considèrent comme satisfaisante.

Admettons maintenant qu'il existe une deuxième cause possible de ce type, par exemple la fréquentation d'un lycée réputé. Il faut alors distinguer au moins quatre sous-populations, caractérisées par les propriétés suivantes : capital culturel élevé et lycée réputé ; capital culturel élevé et lycée non réputé ; capital culturel non élevé et lycée réputé ; capital culturel non élevé et lycée non réputé. « Irène » serait la cause de la mention « Très bien » si et seulement si ce prénom rendait la mention « Très bien » plus probable dans chacune de ces quatre sous-populations *(voir la figure 3)*.

De façon générale, selon cette ligne d'analyse, C cause E si et seulement si C rend E plus probable dans toutes les sous-populations que permettent de définir les causes possibles de E autres que C (et, pour être précis, non causées par C). Cette idée a été introduite en 1979 par la philosophe britannique et américaine Nancy Cartwright. Puisque les sous-populations définies par les causes de E autres que C ne sont pas identiques aux sous-populations définies par les causes de C autres que E, cette analyse n'est pas symétrique. Cette démarche a été précisée dans les années 1980, en particulier par N. Cartwright elle-même. Elle a abouti à une analyse de la causalité que les philosophes considèrent comme satisfaisante.

Toutefois, même si elle constitue un progrès conceptuel, cette analyse de la causalité a ses limites. En effet, il est impossible de transposer directement les théories probabilistes de la causalité au domaine de la méthodologie scientifique : les scientifiques ne peuvent pas les utiliser comme seul guide pour identifier des relations causales à partir des corrélations qu'ils observent dans la nature ou dans leurs expériences. Pourquoi ?

D'abord, ces théories sont circulaires et, en conséquence, il est impossible de même commencer à mener la tâche visant à déterminer si C cause E. En effet, pour déterminer si C cause E, il faudrait selon ces théories déterminer si C augmente la probabilité de E dans toutes les sous-populations définies par les causes de E autres que C. Pour déterminer si C cause E, il faut donc connaître toutes les autres causes de E : en d'autres termes, il faut connaître les causes de E pour déterminer quelles sont les causes de E ! Ainsi, les théories probabilistes de la causalité ne permettent pas à elles seules de commencer à acquérir des connaissances causales.

Ensuite, les théories probabilistes ne définissent pas de principe de clôture *a priori* qui viendrait délimiter l'ensemble des

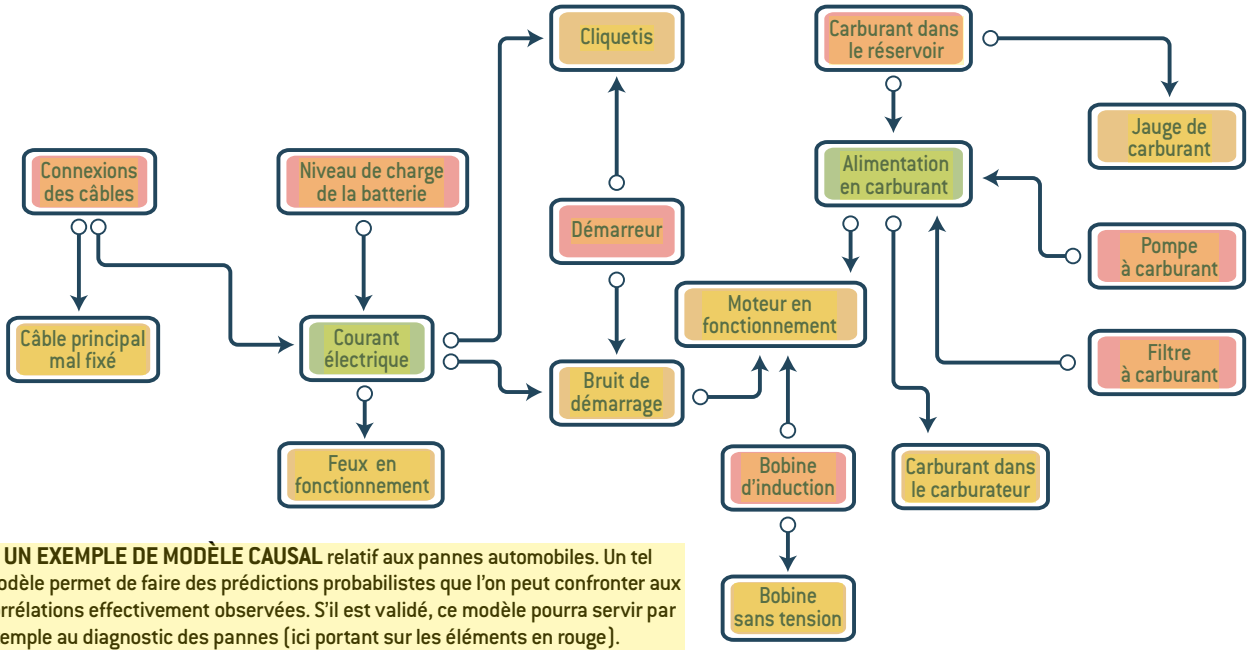
Corrélations sans lien de causalité

En 1956, le philosophe d'origine allemande Hans Reichenbach a caractérisé en termes probabilistes des situations où une cause C a deux effets E et E', ces deux effets n'étant pas liés par causalité (E n'entraîne pas E', et E' n'entraîne pas E).

Cette caractérisation est la suivante :
1. $p(E \mid E') > p(E \mid \text{non-}E')$
2. $p(E \mid C) > p(E \mid \text{non-}C)$
3. $p(E' \mid C) > p(E' \mid \text{non-}C)$
4. $p(E \mid E' \text{ et } C) = p(E \mid C)$
5. $p(E \mid E' \text{ et non-}C) = p(E \mid \text{non-}C)$.

Selon la clause (1), E et E' sont positivement corrélés. Les clauses (2) et (3) indiquent qu'il en va de même de E et C d'une part, et E' et C d'autre part. Jusque-là, il n'y a pas de différence entre E, E' et C.

Mais les clauses (4) et (5) isolent une propriété spécifique à C, qui traduit en termes probabilistes son statut causal particulier dans le système liant les trois propriétés C, E et E'. Selon (4), E et E' ne sont pas corrélés dans la sous-population des individus qui ont la propriété C ; selon (5), ils ne le sont pas non plus dans la sous-population des individus qui n'ont pas la propriété C. Autrement dit, une fois qu'on sait si C est le cas ou non, E' ne nous apprend plus rien à propos de E. Pour décrire cette situation, on dit souvent que C fait écran entre E et E'.



4. UN EXEMPLE DE MODÈLE CAUSAL relatif aux pannes automobiles. Un tel modèle permet de faire des prédictions probabilistes que l'on peut confronter aux corrélations effectivement observées. S'il est validé, ce modèle pourra servir par exemple au diagnostic des pannes (ici portant sur les éléments en rouge).

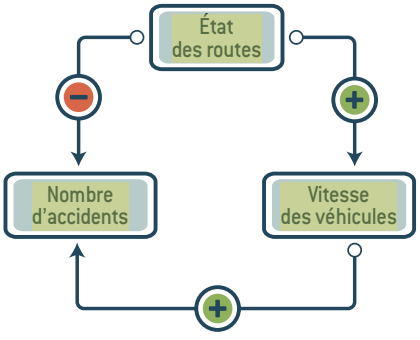
causes possibles d'un effet donné. Elles ne permettent donc jamais de savoir si l'on a identifié toutes les causes d'un effet E, ni par conséquent de savoir si l'on a identifié toutes les causes de E autres que C, ni donc d'établir que C cause E.

En dépit de ces deux obstacles, il existe aujourd'hui au moins trois types de méthodes pour tirer des conclusions causales à partir de corrélations. Comment est-ce possible ?

Une des méthodes : les essais randomisés

Le premier type de méthodes repose sur l'analyse suivante : ce que la circularité et l'absence de principe de clôture interdisent précisément, c'est de tirer des conclusions causales à partir des probabilités dans des contextes où toutes les causes sont présentes en même temps, où personne n'agit sur elles et où le facteur C, dont on cherche à déterminer s'il est une cause de E, a le même statut que les autres causes de E. Or l'expérimentation permet d'éviter cela : la première stratégie consiste ainsi à intervenir sur le système étudié et à s'intéresser non pas à la probabilité de l'effet E quand sa cause supposée C est observée, mais à sa probabilité quand l'expérimentateur impose C.

C'est exactement ce qui se passe dans un essai clinique randomisé, où l'on administre à chaque participant soit le traitement étudié, soit un placebo, selon le résultat d'un tirage au sort. Les probabilités de guérison après une telle intervention permettent de déterminer si le traitement cause la guérison.



5. UNE RELATION DE CAUSALITÉ ne se traduit pas forcément par une corrélation. Ainsi, il existe un lien causal entre l'état des routes et le nombre d'accidents, mais il est possible que la corrélation n'apparaisse pas. Toutes choses égales par ailleurs, un meilleur état des routes réduit le nombre d'accidents ; mais les conducteurs tendent à rouler plus vite si la route est meilleure, ce qui accroît la probabilité d'accidents. Les deux effets peuvent se compenser, auquel cas il n'y aurait pas de corrélation.

Plus précisément, admettons que le tirage au sort garantisse que les différentes sous-populations à considérer soient représentées de façon égale dans les deux groupes ; une probabilité supérieure de guérison dans le groupe ayant reçu le traitement implique que la probabilité de guérison est supérieure dans au moins une de ces situations, et donc que le traitement cause la guérison au sens des théories probabilistes *(voir l'encadré page 60)*. Ce résultat vaut même si l'on n'a pas identifié les causes de guérison autres que le traitement. Il n'est donc pas nécessaire d'identifier toutes les causes de E différentes de C en vue de déterminer si C cause E. Les deux obstacles que nous avons mis au jour, circularité et absence de principe de clôture, sont ainsi contournés.

Une deuxième stratégie émane d'une analyse un peu différente : les obstacles cités plus haut interdisent de tirer des conclusions causales à partir des seules corrélations, et en l'absence de connaissances causales préalables. En ce sens, ils interdisent d'induire des conclusions causales. Mais l'induction n'est pas le seul mode de raisonnement disponible : les méthodes du second type

Figure 4 d'après K. W. Przytula et al., IEEEAC paper 1187, 2002

ont un caractère hypothético-déductif. Qu'est-ce que cela signifie ?

Pour déterminer si le prénom Irène cause la mention « Très bien », un raisonnement hypothético-déductif débuterait par la formulation d'une hypothèse sur les causes de la mention (lesquelles pourraient inclure le capital culturel des parents, la réputation de l'établissement fréquenté, l'âge du candidat, etc.) et sur les relations causales qu'elles entretiennent avec le prénom. Généralement, on nomme « modèle causal » une telle hypothèse, que l'on représente par un graphe : les nœuds de ce graphe correspondent aux phénomènes pris en compte (le prénom, la mention, les causes de la mention…) et les flèches d'un nœud à l'autre représentent des relations de cause à effet *(voir la figure 4)*.

En utilisant les théories probabilistes de la causalité, on tire des conclusions probabilistes d'un tel modèle. Si ces conclusions contredisent les observations, le modèle est rejeté ; si elles sont valides, il peut être considéré comme confirmé. Ici, le problème de la circularité et celui de l'absence de principe de clôture sont résolus dès lors que l'on définit un modèle causal. D'une part, en effet, tout modèle causal porte sur un certain nombre de phénomènes bien déterminés et, de ce fait, il véhicule une réponse à la question de la clôture laissée ouverte par les théories probabilistes. D'autre part, pour chacun des phénomènes qu'il représente, le modèle formule une hypothèse relativement à l'ensemble de ses causes. La circularité des théories probabilistes n'intervient donc plus.

Reste que l'expérimentation et l'hypothético-déduction ont chacune des limites. Un essai randomisé n'est que rarement envisageable – par exemple, il ne l'est pas dans le cas du prénom et de la mention. De même, le succès des approches hypothético-déductives dépend largement des hypothèses qui auront pu être formulées.

Toutefois, des algorithmes développés depuis les années 1990 par les équipes de Judea Pearl, à l'Université de Californie à Los Angeles, et de Clark Glymour, Peter Spirtes et Richard Scheines, à l'Université Carnegie Mellon de Pittsburgh, visent à contourner les obstacles que nous avons mis en évidence d'une façon qui soit à la fois observationnelle et inductive.

Des algorithmes utiles

Ces algorithmes prennent pour entrée les corrélations entre phénomènes d'un ensemble préalablement fixé. Le problème constitué par l'absence de principe de clôture est donc résolu exactement de la même façon que dans le cadre hypothético-déductif : en isolant d'emblée un ensemble fini de phénomènes. Mais ici, aucune hypothèse causale n'est formulée. À partir des seules corrélations – les informations en termes de probabilités –, on construit tous les graphes susceptibles d'être le modèle causal correct pour l'ensemble de phénomènes considéré. L'obstacle de la circularité des théories probabilistes de la causalité est donc contourné.

Ces algorithmes font appel à une analyse un peu différente (et un peu moins satisfaisante) de celle qu'utilisent les théories probabilistes circulaires, avec la conséquence importante que des informations purement probabilistes suffisent à déterminer si deux phénomènes de l'ensemble considéré ont un lien de causalité direct, sans l'intermédiaire d'un autre phénomène.

Cependant, l'analyse qui fonde ces algorithmes admet des contre-exemples, qui ont fait l'objet d'un vif débat dans les années 2000. Notamment, elle suppose

LES ESSAIS CLINIQUES RANDOMISÉS

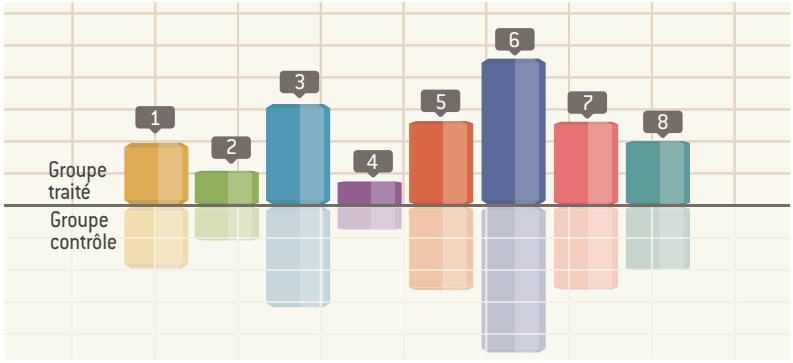
Le premier essai clinique randomisé ayant donné lieu à une publication scientifique a été mené en Grande-Bretagne dans les années 1940*. L'étude portait sur l'effet d'un antibiotique, la streptomycine, sur la tuberculose pulmonaire.

Cent neuf patients, âgés de 15 à 25 ans, y ont participé. Ils ont été répartis en deux groupes, l'un recevant le traitement, l'autre non (le « groupe contrôle »). La répartition a été effectuée selon une procédure aléatoire garantissant que deux patients du même sexe et relevant du même hôpital aient la même probabilité d'appartenir au groupe recevant la streptomycine.

Cette façon de procéder vise à répartir de manière similaire dans les deux groupes les autres facteurs que la streptomycine, le sexe et l'hôpital de prise en charge susceptibles d'affecter causalement l'évolution de l'état de santé des patients.

Imaginons que, dans le cas considéré, il existe trois facteurs de ce type : l'âge, le régime alimentaire et le niveau culturel. Si chacun de ces facteurs est susceptible de deux modalités, huit situations peuvent se présenter. La randomisation (le choix aléatoire des patients à traiter parmi les participants à l'essai clinique) vise à faire en sorte que ces huit situations soient représentées de manière comparable dans le groupe traité et dans le groupe contrôle *(voir la figure)*. Cette procédure est désormais la règle dans les essais cliniques rigoureux, qui doivent par ailleurs porter sur un nombre suffisant de patients pour que la randomisation soit efficace.

*G. Marshall *et al.*, *British Medical Journal*, vol. 2(4582), pp. 769-782, 1948.



EN SUPPOSANT QUE L'ÂGE, le régime alimentaire et le niveau culturel puissent influer sur l'évolution de l'état de santé, et que chacun de ces facteurs ait deux valeurs pertinentes seulement, l'ensemble des patients constitue huit sous-populations. Grâce à la randomisation, les individus de chaque sous-population sont à peu près également répartis entre le groupe traité et le groupe contrôle.