|  |
| --- |
| **[Vers une science de la causalité](http://feedproxy.google.com/~r/internetactu/bcmJ/~3/LOEW74d9RX8/?utm_source=feedburner&utm_medium=email)** - [InternetActu.net](http://www.internetactu.net/)  Posted: 10 Sep 2018 10:00 PM PDT  Depuis l’invention des statistiques, la science ne cherche plus à se prononcer sur la notion de causalité. Corrélation n’étant pas causalité, les statisticiens se refusent à porter un jugement sur quel phénomène « cause » tel autre. La causalité n’est que la forme extrême de la corrélation, autrement dit, ce qui se passe quand une corrélation atteint 100 % de succès.  Pourtant, nous explique [**Judea Pearl**](http://bayes.cs.ucla.edu/jp_home.html) ([**Wikipédia**](https://fr.wikipedia.org/wiki/Judea_Pearl), [**@yudapearl**](https://twitter.com/yudapearl)), dans son livre [***The Book of Why***](https://amzn.to/2Mg2F4A) (Basic Books, 2018 non traduit), écrit en collaboration avec [**Dana Mackenzie**](http://danamackenzie.com/), il nous faut une science de la causalité ! Celle-ci existe déjà, elle s’est développée tout au long du XXe siècle, mais très discrètement. Selon Pearl, elle est un élément indispensable à notre compréhension du monde. Notamment, dans le domaine de l’intelligence artificielle, ce n’est que lorsqu’une parfaite compréhension de la notion de causalité sera atteinte que nous pourrons créer des machines véritablement intelligentes.  Judea Pearl n’est pas un perdreau en matière de recherche : aujourd’hui âgé de 81 ans, c’est son premier ouvrage de vulgarisation. Il est plus coutumier de travaux obscurs consacrés aux hautes sphères de l’IA et des mathématiques, et il a obtenu en 1991 le prix Turing pour ses recherches sur les « [**réseaux bayésiens**](https://fr.wikipedia.org/wiki/R%C3%A9seau_bay%C3%A9sien)« , couramment utilisés aujourd’hui en IA.  **Deux langages pour une nouvelle science**    Comme toutes les disciplines scientifiques, cette nouvelle science de la causalité dispose d’un langage, ou plutôt de deux. Le premier est exclusivement graphique. Il s’agit de diagrammes fléchés qui indiquent l’enchaînement causal supposé. L’autre langage est plus formel, plus proche des maths. Dans son livre, Pearl favorise largement l’usage des diagrammes, plus simples à comprendre, mais donne quelques exemples « d’équations causales ». Celles-ci se caractérisent par l’intervention d’un nouvel opérateur, le « do » (pour « faire » en anglais). Cet opérateur permet de distinguer dans un calcul les modifications observées de celles qui sont le produit d’une intervention volontaire. Autrement dit, observer ce qu’il se passe lorsque les gens mangent des brocolis ne sera pas formulé de la même manière qu’une expérience où un groupe a mangé des brocolis pendant six mois, tandis que le groupe contrôle n’en mangeait jamais.  L’introduction ce petit opérateur change profondément les choses. En fait, une fois qu’on a pris en compte la possibilité d’une intervention, il est possible de calculer ses conséquences sans pour autant la réaliser effectivement : « *L’une des réalisations majeures de la Révolution causale a été d’expliquer comment prédire les effets d’une intervention sans la mettre en pratique. Cela n’aurait jamais été possible si, avant tout, nous n’avions pas défini l’opérateur de manière à pouvoir poser la bonne question et, deuxièmement, imaginé un moyen de l’imiter par des moyens non invasifs. »*  Pour Pearl, l’usage d’un langage formel est nécessaire à la compréhension d’un phénomène. Rappelons qu’il est lui-même chercheur en IA. « *Tout d’abord, dans le monde de l’IA, explique-t-il, vous ne comprenez pas vraiment un sujet avant de pouvoir l’enseigner à un robot. C’est pourquoi vous me verrez mettre l’accent sur la notation, le langage, le vocabulaire et la grammaire (…). C’est incroyable tout ce qu’on peut apprendre en suivant simplement la grammaire des énoncés scientifiques. Mon insistance sur le langage vient aussi de la conviction profonde que la langue façonne nos pensées. Vous ne pouvez pas répondre à une question que vous ne pouvez pas poser et vous ne pouvez pas poser une question pour laquelle vous n’avez pas de mots.*  »    Voici un exemple très simple donné par Pearl d’un tel diagramme causal. Il concerne le fonctionnement d’un peloton d’exécution. Il y a d’abord une décision du tribunal, qui transmet l’ordre au capitaine du peloton, constitué de deux tireurs. Que se passe-t-il si on intervient et que le tireur B décide ne pas tirer ? On supprime la flèche allant de B à l’exécution. Mais cela ne change rien puisque A tire de toute façon. Que se passe-t-il si finalement la cour martiale ne donne pas l’ordre d’exécution ? On efface alors la première flèche. Ce second diagramme est-il possible ? Pour un ordinateur classique, explique Pearl, la question ne se pose pas. La règle est que les soldats ne tirent que si la cour en a donné l’ordre. Mais avec la causalité, explique Pearl, « *nous devons apprendre aux ordinateurs à briser les règles »*. Un des deux soldats peut décider de tirer même contre l’avis de la cour.  On le voit, c’est bête comme chou. Mais vous n’imaginez pas à quel point ces diagrammes deviennent vite complexes tout au long du livre. Si l’ouvrage de Pearl est destiné au grand public, il ne faut pas s’imaginer pour autant qu’il soit simple d’accès. J’avoue qu’il m’a souvent perdu.  Une chose différencie cette approche causale de l’approche statistique classique. Elle implique, de la part du chercheur, une mise en œuvre de sa connaissance du monde. Pour lui : « *(…) dessiner un diagramme de causalité n’est pas un exercice statistique ; c’est un exercice de génétique, d’économie, de psychologie ou de tout autre domaine d’expertise du scientifique. »*  **Les trois niveaux de compréhension du monde**  Que signifie l’avènement de cette nouvelle science de la causalité pour l’IA, qui rappelons-le, est la spécialité de Judea Pearl ?  Ce dernier divise les progrès dans la compréhension du monde en trois étapes.  Le premier stade est l’observation. Le sujet observe son environnement, remarque des régularités et en déduit un certain nombre de conclusions. En d’autres termes, il effectue des corrélations. C’est le stade atteint par les animaux, et selon Pearl, par les actuels systèmes de Deep Learning, qui sont donc loin, selon lui, d’égaler l’intelligence humaine.  Le second stade est celui de l’intelligence humaine, c’est celui de l’intervention (le fameux opérateur « do »). Le sujet effectue une action pour voir ce qui se passe. Il compare les résultats de son acte à la situation antérieure, lorsqu’il n’était pas intervenu. C’est le niveau où on se pose la question « comment… ? »  Le troisième stade, atteint par des humains matures, est celui des « conditions contre-factuelles », autrement dit, on élabore des scénarios imaginaires en se demandant « et si… ? ». Toute interrogation sur les causes d’un phénomène se déroule à ce niveau. Autrement dit, se demander « pourquoi » est en en fait une question contre-factuelle déguisée. Si on revient à l’exemple du diagramme plus haut, on se demande par exemple : « que se passerait-il si la cour ne donnait pas l’ordre d’exécution ? »  Pearl résume ainsi sa théorie : « *le niveau 1 traite du monde tel qu’il est perçu, le 2 d’un monde meilleur, mais néanmoins perceptible, tandis que le niveau 3 traite d’un monde qui ne peut être perçu (parce qu’il contredit ce qui est perçu) ».*  **Les limites des data**  A quel stade se trouvent actuellement les machines intelligentes ? Pour Pearl elles restent coincées au niveau 1, comme d’ailleurs les statistiques traditionnelles : elles ne peuvent répondre à la question « pourquoi ». « *Les données peuvent vous dire que les gens qui ont pris un médicament se sont rétablis plus rapidement que ceux qui ne l’ont pas pris, mais elles ne peuvent pas vous dire pourquoi. Peut-être que ceux qui ont pris le médicament l’ont fait parce qu’ils pouvaient se le permettre et se seraient rétablis aussi rapidement sans lui. »*  En opposition avec sa nouvelle science de la causalité, les data sont incapables d’établir une flèche du temps cohérente. Si A est corrélé à B, B est également corrélé à A. La causalité, elle, ne va que dans un sens.  Dans un monde gouverné uniquement par les data, « *par exemple, les patients éviteraient d’aller chez le médecin pour réduire la probabilité d’être gravement malade ; les villes licencieraient leurs pompiers pour réduire l’incidence des incendies ; les médecins recommanderaient un médicament aux hommes et aux femmes, mais pas aux patients dont le sexe n’a pas été divulgué ; etc. »*  Pearl ne minimise pas les progrès actuels en intelligence artificielle, notamment les technologies du Deep Learning, mais perçoit également leurs limites. « *Les réussites du Deep Learning ont été vraiment remarquables et ont surpris nombre d’entre nous. Néanmoins, le Deep Learning a principalement réussi en montrant que certaines questions ou tâches que nous pensions difficiles ne le sont pas en réalité. »*  Pour tester les futures machines intelligentes, Pearl envisage un mini-test de Turing. Au lieu d’avoir à tromper ses interlocuteurs en se faisant passer pour un humain comme c’est le cas actuellement du [**test de Turing**](https://fr.wikipedia.org/wiki/Test_de_Turing), ce mini-test serait (en apparence) plus facile : « *L’idée est de prendre une histoire simple, de la coder sur une machine, puis de tester si la machine peut répondre correctement aux questions causales auxquelles un être humain peut répondre. »*  L’accès au niveau 3 de la compréhension du monde aurait aussi une autre conséquence. Il permettrait enfin de créer des robots « moraux ». Pour Pearl, un système de règles comme celui des [**trois lois d’Asimov**](https://fr.wikipedia.org/wiki/Trois_lois_de_la_robotique) est une impasse. Il faut permettre au robot de réfléchir à ses propres actions, de se demander « pourquoi » ?  Pourquoi ce lien entre la réflexion sur « des mondes qui n’existent pas » et la conscience morale ? Parce que cela permet de se demander « comment on aurait pu faire autrement ». C’est cette interrogation qui nous permet de réfléchir aux conséquences de nos actes. Si nous arrivons à créer une telle IA, insiste Pearl (et il affirme que certains de ses étudiants mettent actuellement au point des algorithmes allant dans ce sens), alors il n’y aura aucune raison d’avoir peur de l’intelligence des machines.  Rémi Sussan |