

La virologie informatique : genèse d'un concept ?

Anne Bonfante¹ et Jean-Yves Marion²

Le terme de « virus informatique » a été employé pour la première fois par Cohen dans sa thèse, en 1986. Il lui avait d'ailleurs été suggéré par son directeur de thèse, Adleman. En effet, dès l'introduction du texte de sa thèse, il écrivait : « On pourrait définir un virus approximativement comme une séquence de symboles qui, selon une interprétation et dans un environnement donné, provoque la modification d'autres séquences de symboles dans ce même environnement, de telle sorte qu'elles contiennent des virus (eux-mêmes éventuellement modifiés) ».

Le choix du terme « virus » n'est pas anodin, ni neutre sur le plan scientifique. En effet, il implique d'emblée une analogie entre le monde de l'informatique et celui de la biologie. On peut penser que ce choix a été indiqué par le fait que la nature d'un virus biologique correspondait bien à celle d'un virus informatique « infectant » des programmes (encore un emprunt à la biologie!). En d'autres termes, la biologie fournit une bonne représentation de ce qui se passe dans le monde informatique. C'est à ce sujet que nous voudrions nous interroger : comment le choix d'un terme induit un type de démarche scientifique, comment un concept circule d'une discipline à l'autre, et s'il suffit à définir un domaine de recherche.

L'analogie avec la biologie est dictée par plusieurs correspondances : la structure d'un virus, sa fonction et son mode de reproduction. La structure d'abord : un virus biologique est composé d'un filament d'acide nucléique qui correspond à une séquence de lettres, comme un programme. La fonction des deux types de virus ensuite : ils induisent une infection, et sont capables de muter. Le mode de reproduction enfin : les virus informatiques semblent avoir en commun avec leurs « collègues » biologiques, de pouvoir se recopier indéfiniment. Le mécanisme de leur réplication semble fonctionner d'une façon similaire au monde du vivant, ainsi que les travaux de von Neumann sur les automates auto-reproducteurs l'ont montré (dès 1966). Pour mener l'analogie plus loin, on peut se rappeler le mot du prix Nobel Lwoff : « un virus est un virus », ce qui est une façon de définir par auto-référence, un virus biologique. Or, un virus informatique est un programme auto-reproducteur, dont la conception théorique est liée au second théorème de récursion de Kleene, qui est un exemple de solution d'équations auto-référentes. L'analogie avec la biologie recouvre donc plusieurs dimensions épistémologiques.

Pour nous donner un autre éclairage et une perspective historique et philosophique, on peut faire un petit détour par l'oeuvre de Descartes, qui est à compter comme une étape décisive dans la naissance de la science moderne.

Descartes au XVII^{ème} siècle, avait suggéré que pour mieux comprendre le fonctionnement du corps, on pouvait se le représenter comme un mécanisme (une horloge par exemple). Quand on ne sait pas exactement comment analyser un phénomène, il peut être utile de s'aider d'un modèle, qui permet au moins de se représenter son fonctionnement. Ainsi pour le corps ou pour le comportement des animaux : il est éclairant de se le représenter comme étant semblable à celui d'une machine. Ce qui compte, c'est de pouvoir se donner une image du fonctionnement des choses, quand justement elle ne nous tombent pas sous le sens ou ne sont pas directement observables : la représentation permet de faire des hypothèses, qu'on pourra peut-être vérifier par l'expérience plus tard. Mais bien entendu, un modèle épistémologique comme celui de la machine n'a de valeur que par la compréhension scientifique qu'il permet de développer : il ne s'agit nullement de donner à ce

¹ Université de Bourgogne. Email : anne.bonfante@libertysurf

² Ecole nationale supérieure des Mines de Nancy, Loria-INPL.
Email : Jean-Yves.Marion@loria.fr

modèle une valeur ontologique (en d'autres termes, il ne s'agit pas de prendre ce modèle pour la réalité), mais seulement épistémologique.

Ensuite, cette objection montre que, parfois, la représentation elle-même intéresse plus, ou frappe davantage l'imagination, que la réalité qu'elle est censée représenter (ainsi pour les animaux et les machines). Le danger est alors que la représentation prenne davantage de place que la réalité qu'elle est censée permettre d'expliquer. Quel rapport avec la virologie informatique ? D'abord, et peut-être n'est-ce qu'une anecdote, si les mauvais lecteurs de Descartes pensent que celui-ci a « dit que les animaux étaient des machines », et lui attribuent très à tort la confusion entre êtres vivants et automates, d'autres, qui sont leur reflet, font la confusion inverse, et n'hésitent pas à nous proposer des « vaccins » pour nos ordinateurs, et des antivirus qui affichent un stéthoscope, une gélule ou un caducée comme argument publicitaire ... Une autre erreur serait d'oublier la fonction représentative de la biologie, et de ne plus distinguer l'hypothèse qui permet de travailler, de la réalité sur laquelle on travaille.

Le danger d'une analogie est de réduire deux objets différents à un seul, sans tenir compte des particularités de chaque monde. Ainsi, si la biologie est un domaine d'inspiration légitime pour l'informatique (pour la virologie en particulier), elle n'épuise pas la multiplicité des phénomènes du monde informatique. On peut objecter par exemple que les infections informatiques s'apparentent bien plus souvent à des armes de guerre, et que le lexique du conflit est tout autant utilisé pour décrire les virus : quelle est alors la bonne représentation ? La biologie ou la guerre ?

Il est vraisemblable que la présence de deux (ou plusieurs) domaines de représentation à la fois montre surtout que le monde de l'informatique offre une véritable richesse conceptuelle, puisque un seul type d'analogie n'épuise pas les perspectives de recherche. La valeur de la représentation qu'offre la biologie au monde informatique est évidente. Par ailleurs, il est significatif que l'informatique est dorénavant elle-même une source d'inspiration pour les recherches en biologie (notamment à propos du fonctionnement des cellules).

Bibliographie

- Adleman, Leonard M. An abstract theory of computer viruses. In *Advances in Cryptology*, LNCS 403, 1988.
- Cohen, Fred. *Computer Viruses*. PhD thesis, University of Southern California, January 1986.
- Descartes, René. *Le monde et le traité de l'homme*. Edition Alquié, Classiques Garnier, 1988.
- Filiol, Eric. *Les virus informatiques : théorie, pratique et applications*. (Ed. Springer, 2004)
- Kleene, Stephan C. *Introduction to Metamathematics*. Van Nostrand, 1952.
- von Neumann., *John Theory of Self-Reproducing Automata*. University of Illinois Press, Urbana, Illinois, 1966.