Le complexe de Frankenstein revisité

Stéphane Natkin[1]

1. CNAM /CEDRIC
292 rue St Martin 75003 Paris
stephane.natkin@cnam.fr

Résumé. Dans ce court texte nous analysons, par le biais des fameuses trois lois de la robotique, l'importance des dangers liés à la numérisation de notre société et la perception que nous en avons. Après avoir tenté de montrer que, d'un point de vue scientifique et technologique de tels risques existent, nous concluons en analysant les rapports d'usages qui sont sans doute au cœur de cette problématique.

Abstract. In this paper we analyze, through the famous three laws of robotics, the importance of threats and dangers related to the digitalization of our society and the way they are perceived. We show first that, from a scientific and technical point of view, the fear of these dangers is justified. Our conclusion tries to show that the core of this problematic relies on the practice ad usability of the "new" digital world.

Mots-clés : societe numérique, sécurité, usage.

Keywords : digtal society, safety, security, usability

1. Les ordinateurs fous[1]

Le début des romans "le cycle des robots" d'Isaac Asimov (Asimov 2001) a pour héroïne Suzanne Calvin, une robot- psychologue, dont l'activité consiste à traiter la folie de ces machines très perfectionnées. L'action se situe dans un futur ou les robots sont devenus indispensables à toute activité humaine. Une erreur comportementale d'un robot (résultant d'une faiblesse psychologique ou de l'influence de la méchanceté humaine) peut avoir des conséquences catastrophiques. Les robots sont régis par les fameuses trois lois de la robotique qui sont censées protéger les hommes (Robotique 2058) :

- Un robot ne peut porter atteint à un être humain ni, restant passif, laisser cet être humain exposé au danger
- Un robot doit obéir aux ordres donnés par des êtres humains sauf quand de tels ordres sont en contradiction avec la première loi,

[1] Ce texte est une adaptation de l'introduction de (Natkin 2002)

- Un robot doit protéger sa propre existence dans la mesure ou une telle protection ne s ‘oppose pas à la première et seconde loi.

Mais la puissance des robots a pour conséquence d'exacerber la peur ancestrale des machines, dénommée par Asimov "complexe de Frankenstein".

Les robots restent à inventer mais nous avons déjà peur de nos ordinateurs. Cette peur s’est longtemps concentrée sur les effets possibles d’une erreur de programmation dans le contrôle de processus critique (Laprie 95). Elle se porte à présent sur le détournement possible des nouvelles technologies de l’information et la communication par soit des pirates et autres saboteurs sans vergogne, soit par un état aspirant au meilleur des mondes (Erickson 2008) (Gollmann 2011). Souffrons-nous du complexe de Frankenstein ou faut-il craindre avec raison les effets pervers d’une informatisation trop rapide de la société ?

Pour traiter ce problème il faut d’abord comprendre que les risques résultant d’une erreur humaine et ceux découlant d’un usage volontairement dangereux des ordinateurs sont très liés. Ces risques résultent essentiellement, dans les deux cas, d’une mauvaise spécification : nous ne savons pas définir exactement ce que nos machines doivent faire ou ne pas faire. Nous ne savons pas encore implanter les trois lois de la robotique qui limiteraient drastiquement leurs comportements dangereux et nous confions chaque jour des opérations de plus en plus complexes à nos systèmes informatiques.

Remarquons ensuite que cette peur a des fondements raisonnables. Les ordinateurs peuvent devenir fous et cette folie peut prendre de multiples formes. La dépression électronique la plus courante s'exprime par un refus clair et définitif de faire quoi que ce soit, mais il existe des pathologies plus complexes et plus rares. Par exemple l'ordinateur peut faire trop tôt ou trop tard ce qu'il devait faire ou, cas de folie grave, accomplir des actions qui sont fort différentes de celles attendues. Par ailleurs votre ordinateur a un caractère influençable. Il peut se laisser pervertir par un pirate et détruire votre courrier, transformer votre écran en une œuvre d’art minimaliste ou inonder la planète de messages pornographiques. Il peut se soumettre à un marchand pour guetter à votre insu vos comportements de consommateurs. Il peut devenir un agent d’un état policier et surveiller vos communications Le caractère imparfait de ces faibles machines est bien connu puisque "c'est la faute à l'ordinateur" constitue une excuse de plus en plus utilisée aux défaillances humaines. Les conséquences des dysfonctionnements sont dans la majorité des cas assez bénignes : Tout utilisateur d'un ordinateur personnel a eu la joie de se "retaper" deux ou trois fois la même chose, suite à la "perte d'un fichier". Toutefois, comme les ordinateurs se mêlent chaque jour un peu plus à toutes les activités humaines, il existe des domaines d'utilisation de l'informatique où les états d'âme de nos collaborateurs électroniques peuvent avoir des conséquences considérables. Dans le domaine des cataclysmes actuellement observables ou observés, citons la paralysie des serveurs Web, la disparition de sommes considérables, la faillite d'une entreprise qui ne peut plus facturer, la création d'embouteillages monstrueux, l'échec du tir d'Ariane ou une panne de courant paralysant une métropole. Le contrôle informatisé de processus techniques de plus en plus complexes augmente le risque de

catastrophe. Le développement d'applications qui reposent sur l'authentification numérique (comme la signature électronique) crée une dépendance dont les effets individuels ou collectifs peuvent être désastreux. Le jour où les systèmes électroniques ne sauront plus reconnaître votre carte d'identité à puce, existerez vous encore ? Constatons à nouveau que ces évènements peuvent aussi bien résulter d'une défaillance ou d'un sabotage.

La robot-psychologue, Suzanne Calvin, vieille fille frustrée, a beaucoup plus d'affinité pour les robots que pour les hommes et utilise sa grande intelligence à laver les soupçons qu'à chaque incident l'humanité fait peser sur ses protégés. Les nouvelles écrites par Asimov permettent à Susan Calvin de montrer que les robots ne sont en général que les victimes de l'ignorance, la bêtise ou la malignité humaine. Les statistiques montrent que ce constat peut être appliqué à nos ordinateurs. La part des sinistres qui peut être imputés à un phénomène accidentel est inférieure à 25%, toutes les autres causes sont des conséquences directes ou indirectes d'un comportement erroné ou malveillant des hommes. Ces statistiques sont sans doutes à prendre avec de grandes précautions, mais elles sont aussi un début de preuve de l'innocence des ordinateurs.

2. La complexité des systèmes d'information et des systèmes informatiques

Comme toute forme de folie, celle des ordinateurs est une tare honteuse que les concepteurs de systèmes informatiques cachent le plus longtemps possible. Lorsque la catastrophe est patente, il est de bon ton de trouver un responsable, montrant l'erreur et par la même la capacité à l'éviter. Ce qui sous-entend, comme dans toute analyse primaire de la folie, qu'il existe un comportement sain et que la déviation à ce comportement peut être guérie soit en punissant les coupables humains, soit en "réparant" le système, soit en l'envoyant à la casse (la nécessité d'un asile pour ordinateurs défectueux n'est pas envisagée, même dans les livres d'Asimov).

Ce raisonnement ne résiste pas à une analyse approfondie. Pour s'en convaincre, il faut chercher à définir en quoi consiste la panne d'un ordinateur telle qu'elle est perçue par ses utilisateurs. Il n'existe d'autre définition que la déviation par rapport à un comportement attendu, recréant ainsi une notion de normalité. Mais, les hommes, créateurs de la machine, ne sont pas supposés déduire par analyse psychologique, sociologique ou éthique, les critères normatifs applicables. Ils devraient, deus ex machina, spécifier ce comportement avant que la machine ne soit créée, ceci permettant de vérifier pendant le processus de conception ou lors des étapes d'essais et de test que le comportement réalisé est conforme à celui attendu.

En toute généralité, cette tâche est, et restera probablement pour fort longtemps, impossible à réaliser. En effet un ordinateur évolue et réagit en fonction du reste de l'univers, composé entre autre de ses utilisateurs, des phénomènes physiques divers qui peuvent l'affecter lui-même ou le système (jeux vidéo, comptabilité, train, centrale nucléaire) qu'il contrôle. Par ailleurs dans un environnement agressif certaines défaillances, et certains comportements anormaux ou malveillants des hommes doivent donner lieu à une réaction contrôlée. Spécifier les comportements

"normaux" de l'ordinateur suppose donc de pouvoir définir toutes les actions, séquentielles ou imbriquées, qui peuvent affecter l'ordinateur et son environnement. Il faut ensuite, pour chaque cas recensé, décrire les réactions attendues de la machine. De plus cette description se doit d'être complète et non ambiguë, donc doit être exprimée selon un système formel. En d'autres termes il s'agit de formaliser une portion de l'univers, celle qui se trouve dans un voisinage proche de la machine. Considérons par exemple le pilotage automatique d'une ligne de métro, dans lesquels les ordinateurs remplacent les conducteurs. Outre les métros, les voies ferrées, avec de nombreux aiguillages, des signaux complexes, un tel système comporte des opérateurs humains qui en général suivent les procédures d'exploitation mais qui peuvent se tromper, des passagers qui parfois empêchent les porte de se fermer pour décoincer leur sac à dos ou qui se jettent sous le métro par désespoir sentimental, des rochers qui peuvent éventuellement tomber sur la voie, des règlements nationaux inadaptés qui définissent diverses limites d'utilisation, des terroristes en mal de sabotage et des lutins mystérieux qui peuvent provoquer des pannes variées de tout ce qui vient d'être mentionné. Le problème est simple : pour définir ce que doivent faire les ordinateurs "normaux", il faut, au préalable avoir mis tout ce petit monde sous forme d'équations. Il existe peut être une forme d'intelligence supérieure qui sait faire cela, mais c'est hors de la connaissance scientifique. Il est probable que le jour ou les hommes sauront résoudre ce problème (à supposer qu'il sache le faire un jour), la plus grande partie des autres problèmes ouverts en matière de sciences physiques et humaines seront clos.

En supposant possible la capacité à spécifier les objectifs attendus, la conception de systèmes informatiques sûrs nécessite de résoudre une seconde classe de problèmes qui est liée à leur complexité intrinsèque. Un ordinateur, tel que celui qui est utilisé pour taper ce texte est le résultat d'un processus de conception réalisé en plusieurs phases. Il comporte la conception du matériel et en amont de ses composants (microprocesseurs, mémoires…), d'un système d'exploitation, puis d'un ou plusieurs programmes comme le traitement de texte. Chacun de ces composants repose essentiellement sur le raisonnement de son concepteur exprimé dans un langage formel. L'ensemble devrait être parfaitement apte à permettre de saisir ce texte. Malheureusement, pour les raisons exposées plus haut, nul ne peut prédire toutes les façons d'utiliser le traitement de texte. Le concepteur du système d'exploitation ne peut connaître tous les usages du système par les logiciels d'applications. Les architectes matériels ont conçus processeurs, mémoires et cartes en pensant à certaines fonctionnalités des futurs systèmes d'exploitation et en ignorant d'autres. Cet ordinateur est donc basé sur la collaboration entre trois (groupes d') humains qui ne peuvent pas se comprendre complètement. Pour comprendre à posteriori comment marche cette machine, il est nécessaire de faire appel à des techniques expérimentales et d'analyses probabilistes, comme pour comprendre l'électeur moyen ou le virus de la grippe.

Les problèmes et risques liés aux usages actuels de l'Internet résultent du même phénomène à l'échelle de la planète. Il fut créé comme une plate-forme d'expérimentation des recherches en matière de télécommunication numérique et un moyen d'échange entre chercheurs universitaires. La technologie de l'Internet et son

déploiement résulte d'un processus complexe, non contrôlé par les administrations, basé sur le volontariat et le sponsoring des universités et de certains constructeurs d'ordinateur. Les principaux objectifs visés étaient la facilité de rattachement au réseau (il suffit d'avoir le téléphone) et la possibilité de développer et d'échanger de nouveaux logiciels expérimentaux. L'Internet répond parfaitement à ces besoins. Par contre, il n'a à priori aucune qualité pour faire du de la diffusion de media, du commerce électronique, de la domotique ou de la communication d'entreprise. L'usage a changé, la spécification des besoins n'est pas encore bien posée. Il ne faut dons pas s'étonner de rencontrer quelques problèmes.

Il résulte donc de ce qui vient d'être énoncé, que, même si ce n'est pas leur faute, il faut avoir peur des ordinateurs. Nous allons nuancer ce jugement en considérant le processus d'apprentissage social qui détermine nos relations avec numérique.

3. De la pratique sociale des ordinateurs

La relation entre les ordinateurs et leurs utilisateurs est devenu un des champs les plus important de recherche et de développement de l'informatique. Aujourd'hui plus de 50% de la puissance informatique est consacrée à l'interaction homme/machine et ce taux ne fait qu'augmenter. Rien de plus normal : presque tous les objets que nous fréquentons et que nous utilisons sont contrôlés par des processeurs, de notre voiture à la porte d'entrée du bureau. Demain ce sera le frigidaire, le masseur intégré dans le fauteuil, le mobilier urbain.... Demain la vie sera numérique, l'Internet des objets et l'intelligence ambiante nous cerneront. Mais la ou la technologie plane, la sociologie et la psychologie rame et notre pratique personnelle nous coule, surtout lorsque l'on n'est pas un « digital native ».

Considérons l'usage du courrier électronique. Le courrier électronique est très vulnérable à toute forme d'attaque : il est possible de le détruire, de le détourner, de modifier l'expéditeur, le destinataire et le contenu sans trop de difficulté. Pourtant cette caractéristique nous échappe en général et nous faisons des « reply to all » généreux et inconsidérés aux messages reçus. Pourtant lorsque nous utilisons le courrier postal "normal" et en voie de disparition, nous nous posons peu de questions : s'il s'agit d'un message d'usage courant nous collons un timbre, s'il a une importance contractuelle limitée nous utilisons le recommandé avec accusé de réception, pour les courriers très importants nous utilisons un porteur ou un huissier. Nous avons une pratique sociale qui nous fait adapter notre usage au risque encouru sans pour autant mener à chaque fois une étude de risques. Il existe des formes équivalentes du courrier électronique, mais nous ne savons pas encore les pratiquer. D'ailleurs, le même flou existe en matière de formule de politesse : selon l'origine de l'expéditeur, un courrier électronique commence par Monsieur, Bonjour, Salut, :-) ou une absence totale de (formule de) politesse. Cet exemple est démultiplié par cent si l'on considère l'ensemble des formes de communications que nous offre la vie numérique actuelle : les SMS, Facebook, le tweet et autres chats...

De façon plus générale, il n'y a pas de pratique sociale définie pour la plus grande partie de notre vie numérique. Nous singeons dans les livre et journaux

électronique les techniques d'édition du papier alors qu'il est évident que ce n'est pas la forme d'écriture et de lecture adaptée (Ciment & Natkin 2011). Nous essayons de protéger le droit des artistes et des créateurs en nous appuyant sur un droit sur la propriété intellectuelle et des règlements des sociétés d'auteurs créés en 1945.

Si les états et les utilisateurs moyens sont mal connectés, les attaquants eux se sont adaptés et démontrent chaque jour que nos mécaniques et nos lois peuvent être contournées que ce soit pour notre bien ou pour notre mal. Les attaques du DNS par déni de services que nous envisagions en 2002 se sont produites (Natkin 2002). Les hackers sont sortis d'une quête à la notoriété pour un militantisme politique dont les Anonymes sont le modèle. Les pirates commencent à gagner de l'argent, ne serait ce qu'en vendant de la musique qui ne leur appartient pas.

Une anecdote démontre les contradictions entre nos pratiques sociales et économique, les mécanismes légaux et le mode numérique (Natkin 2006). Elle concerne les jeux en ligne massivement multi-joueurs (MMOG). Dans ces univers virtuels qui disposent d'une monnaie et d'un système commercial, il n'y a pas de banque centrale. Un logiciel de régulation contrôle les émissions et destructions de monnaie. En 2005 un économiste étudie le système monétaire du jeu « Starwars on Line (SOL) » et constate que la quantité de monnaie détruite par le jeu est supérieur à celle créée. Un phénomène de déflation devrait en résulter or il n'en est rien. Poursuivant son enquête notre économiste découvre que des hackers produisent de la fausse monnaie de SOL. C'est cette monnaie, non identifiable par le régulateur, qui équilibre les comptes. Mais la partie la plus étonnante de l'histoire est que la monnaie SOL pouvait être vendue et achetée sur ebay contre des dollars, certes électroniques, mais générés par le trésor US. Conclusion : si vous désirez vous lancer dans la confection de fausse monnaie il existe deux solutions. La traditionnelle consiste à recruter un graveur, trouver du papier adéquat et une imprimerie cachée. Si vous vous faite prendre vous risquez quelques dizaines d'années de prison. La méthode numérique consiste à fabrique des faux sols électroniques (ou tout autre bien virtuel commercialisable) dans un jeu MMOG ou dans un réseau social puis de les revendre en ligne contre de bons dollars. Si vous êtes pris vous serez interdit de jouer à SOL pour avoir violé le contrat d'achat...

De la même façon, la peur du pilote automatique de métro est en grande partie liée à un usage professionnel non maîtrisé. Dans la plus grande partie des systèmes actuels, cent fois plus de précautions sont prises dans la conception du pilote automatique que dans celle des freins. Pourtant, sans freins, un pilote, automatique ou humain, ne peut rien faire. La différence tient au fait que notre pratique de la technologie des freins à plus de cent cinquante ans et celle des ordinateurs une cinquantaine d'années.

Au fur et à mesure qu'une pratique sociale s'impose, la spécification des besoins peut être déterminée, le niveau des risques acceptables évalué et les moyens techniques nécessaires développés et mis en œuvre. On peut alors envisager de légiférer ou de règlementer. Dans certains cas, il est urgent d'attendre. Il y a une dizaine d'années, de peur de rater un train technologique, la France et la

communauté européenne légiféraient à tour de lois et règlements sur l'usage de systèmes très complexes d'infrastructures à clef publiques, destinés à assurer la sécurité des échanges électroniques. Ces systèmes se sont révélés généralement inadaptés. Ils induisent à la fois un risque de contrôle étatique et celui d'une perte, par les mêmes états, de pouvoirs (comme la délivrance de papiers d'identité) essentiels à un fonctionnement cohérent d'une société démocratique. Lors de la première rédaction de ce texte, le commerce électronique avec le consommateur final commençait à peine à émerger et entre société était encore un concept. On imaginait des solutions de très haute sécurité. Quinze ans plus tard l'administration fiscale a renoncée aux certificats et la banque se contente de mots de passe. L'expérience a permis de définir un équilibre entre sécurité et une facilité d'usage adaptée.

Il existe donc une probabilité raisonnable de pouvoir cohabiter et même collaborer avec les ordinateurs. Il suffit de prendre le temps de savoir ce que nous voulons en faire et comment. Lorsque le problème est bien posé, les solutions techniques, existent déjà souvent et, dans le cas contraire, seront inventées.

Bibliographie

Asimov I. (2001). *Les robots*, J'ai Lu, 2001, Paris

Ciment G., Natkin S. (2011). Quelques media en voie de disparition : de la presse à la bande dessinée, *Impertinence 2011*, La documentation Française, Paris

Erikson J. (2008). *Hacking, the Art of Exploitation,* No Starch Press, californie USA

Gollmann D. (2011). Computer Security, J. Wiley and Sons, NY USA

Laprie J.C. et als (2995). *Le guide de la sûreté de fonctionnement,* CEPADUES, Tolouse

Natkin S. (2002)**.***Les Protocoles de sécurité de l'Internet* - Dunod, 2002

Natkin S. (2006). *Games and Interactive Media: A Glimpse to New Digital Entertainment*, AK Peters, Boston

Robotique. (2058). Manuel de la robotique, 58 édition, USA 2058, cité dans (Asimov 2001)

Biographie

Stéphane Natkin est professeur au CNAM, directeur de l'Ecole Nationale des Jeux et Media Interactifs (ENJMIN) et responsable de l'équipe Media Interaction et Mobilité au CEDRIC. Il est membre du CA du CNAM, du CA et du BE de Cap Digital. Il est l'auteur de nombreuses publication dans le domaine des systèmes critiques, des media interactifs et des jeux vidéos.