

Machines, prédictions et fin du monde

Les paradoxes de la prédiction prouvent-ils l'existence du libre arbitre, et l'imminence de la fin de l'humanité?

Nous sommes déconcertés par les progrès de la science, et il nous arrive de nous réjouir des limitations absolues de l'approche rationnelle du monde. On connaît, par exemple, l'utilisation du non-déterminisme et de la non-localité de la mécanique quantique pour argumenter en faveur du libre arbitre humain ou de la transmission de pensée. On connaît aussi le grand pouvoir de séduction des théorèmes d'incomplétude de Gödel : il n'est pas rare qu'ils soient évoqués à propos de questions étrangères aux mathématiques.

Moins connues, certaines difficultés de la théorie des probabilités et de la prédiction pourraient être exploitées par tous ceux qui ont le souci d'établir la supériorité des humains sur les ordinateurs et les machines. Par jeu, et peut-être aussi pour encourager le renouvellement des discussions entre les mécanistes qui défendent que le cerveau n'est qu'une machine compliquée (et dont les arguments ne sont pas tous mauvais) et les antimécanistes (dont les arguments ne sont pas tous bons), nous allons présenter quelques-unes des difficultés de la prédiction. Ces difficultés montrent que, pour des raisons plus ou moins bien éclaircies, la prédiction est parfois impossible. Nous terminerons par un paradoxe récemment développé par le philosophe canadien John Leslie qui, si on le prenait au sérieux, devrait nous inquiéter gravement.

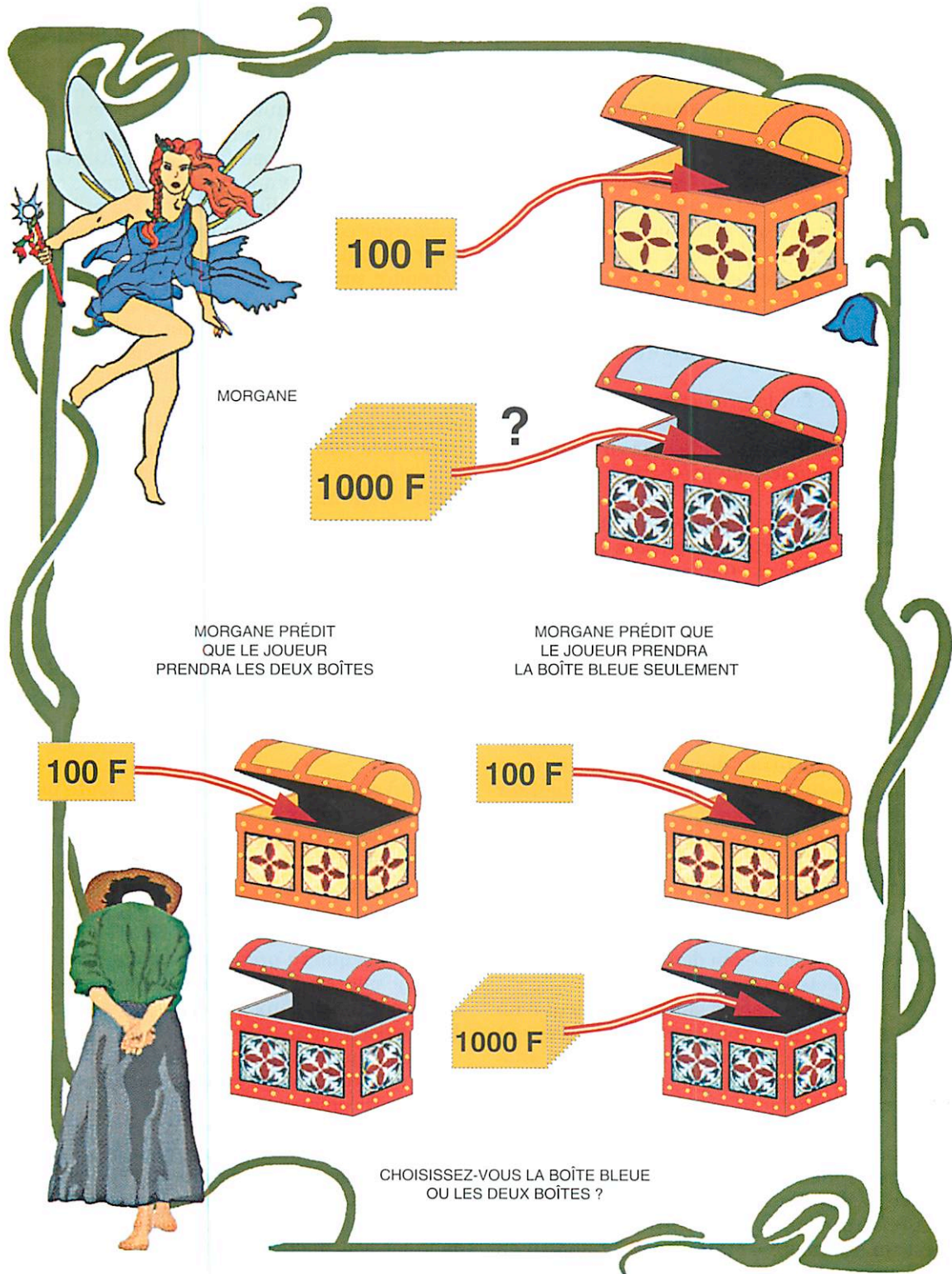
L'ordinateur et l'apéritif

Commençons par le plus simple et le plus spectaculaire des paradoxes de la prédiction, qu'on pourrait appeler le paradoxe de l'apéritif.

Un ami qui vous reçoit chez lui vous offre le choix entre du whisky et du jus d'orange. Vous lui annoncez alors : «Je vais te prouver que je suis un être libre dont tu ne pourras pas prévoir le comportement. Dis-moi ce que tu crois que je vais choisir, et tu verras que tu te trompes.» Si votre ami vous annonce : «Tu vas choisir le whisky», vous lui répondez : «Tu as perdu, je choisis le jus d'orange» ; et s'il vous annonce : «Tu vas choisir le jus d'orange», vous lui répondez : «Perdu! Je choisis le whisky.» Vous pouvez ajouter : «Tu vois, je t'ai prouvé que je suis libre, mais si tu penses qu'une seule expérience ne suffit pas, recommençons.»

Bien sûr, personne n'est dupe de ce genre de farce – sauf peut-être les adolescents dans leurs relations avec leurs parents – et il faudrait beaucoup de naïveté pour croire qu'on établit ainsi la réalité du libre arbitre ou l'irréductibilité du cerveau humain à une machine. Il y a d'ailleurs un moyen radical de montrer qu'on n'établit pas vraiment la preuve du libre arbitre humain avec le paradoxe de l'apéritif : c'est le programme d'ordinateur suivant de cinq lignes qui peut vous remplacer :

1. La fée annonce : (1) Je ne me trompe jamais dans mes prédictions ; (2) Voici une boîte rouge et une boîte bleue ; (3) Tu pourras ouvrir soit la boîte bleue, soit la boîte rouge et la boîte bleue, et tu prendras alors le contenu des boîtes ouvertes ; (4) Dans la boîte rouge, je mettrai 100 francs ; (5) Avant de remplir la boîte bleue, je ferai une prédiction (que je ne te communiquerai pas) : si je prédis que tu prendras la boîte bleue toute seule, je mettrai 1 000 francs à l'intérieur ; si je prédis que tu prendras les deux boîtes, je ne mettrai rien dans la boîte bleue. Que faut-il faire? Deux raisonnements apparemment rationnels conduisent à deux conclusions opposées.



Écrire à l'écran «Que crois-tu que je vais choisir?»

Lire la réponse, la mettre dans la variable *R*,
Si *R* = «Le whisky» alors écrire «Je choisis le jus d'orange»,

Si *R* = «Le jus d'orange» alors écrire «je choisis le whisky»,

Écrire «Tu vois que je suis libre».

Je maintiens, que malgré sa banalité, le paradoxe de l'apéritif est intéressant, car, de la façon la plus élémentaire qui soit, il montre que, dans certaines situations simples et sans mystères, une partie d'un système ne peut prédire un événement futur du système, et cela même si le système global est fini, déterministe et n'est constitué que d'un petit nombre de composants mécaniques.

Si celui qui fait la prédiction n'est pas obligé de la divulguer, il n'y a plus de paradoxe. C'est donc l'appartenance du prédicteur au système et l'obligation d'énoncer sa prédiction qui engendrent ici la difficulté.

N'oublions pas que, à l'inverse, la prédiction est possible dans certains cas assez complexes : un ordinateur isolé est toujours prévisible. Pour connaître le résultat de son calcul lorsqu'on le fera fonctionner, il suffit de connaître son état initial détaillé et de disposer d'une machine analogue qu'on placera dans le même état initial et à qui on fera réaliser le même calcul. Tout est parfaitement fixé à l'instant du début du calcul de l'ordinateur et, tout étant parfaitement déterministe dans l'exécution des programmes, il en résulte qu'aucune incertitude n'entache leur déroulement, même compliqué. Notons d'ailleurs que l'un des buts principaux des constructeurs et des concepteurs d'ordinateurs est d'éviter tout indéterminisme, ce qui parfois n'est obtenu (en particulier pour les mémoires vives) qu'en utilisant des techniques spéciales – comme les codes correcteurs d'erreurs – qui détectent et corrigent tout écart au déterminisme strict.

Par nature donc, le résultat du calcul d'un ordinateur isolé peut toujours être simulé, soit sur une autre machine identique – ce qui simplifie le travail –, soit – et en informatique c'est une opération courante – sur une machine différente qui imitera – on dit parfois qui émulerà – la première.

Dans l'énoncé ci-dessus, il est essentiel de préciser «isolé». En effet, un ordinateur isolé ne peut prendre qu'un nombre fini d'états différents – c'est un automate d'états finis – et donc il en arrive nécessairement à tourner en rond, ce qui rend prévisible non seulement son comportement

dans la minute qui vient, mais aussi son comportement à l'infini. Un ordinateur non isolé qui dispose de mémoires de plus en plus volumineuses à sa demande – comme une machine de Turing – est simulable comme précédemment et donc est prévisible pour la minute qui vient, mais n'est pas prévisible à l'infini. On ne peut pas, par exemple, déterminer par analyse de son état initial et de ses programmes s'il va s'arrêter ou non : c'est la fameuse indécidabilité de l'arrêt. Par ailleurs, un ordinateur connecté à un réseau est imprévisible pour la raison toute simple qu'il peut faire dépendre son comportement futur des informations qu'il échangera par le réseau.

Gardons à l'esprit les deux évidences que nous venons de voir, car elles nous serviront de repère quand, plus loin, la tête nous tournera : (a) pour des raisons immédiates et n'ayant rien à voir avec le libre arbitre, la prédiction est parfois impossible ; (b) il y a des cas non triviaux où la prédiction est possible, comme celui du comportement des ordinateurs isolés.

L'intelligence artificielle mise en doute?

La deuxième situation où la prédiction entraîne des difficultés graves est connue sous le nom de paradoxe de Newcomb. Il fut inventé par le physicien William Newcomb vers 1960 et a été l'objet de discussions acharnées depuis.

Une fée vous dit :

(1) Je ne me trompe jamais dans mes prédictions ;

(2) Voici une boîte rouge et une boîte bleue ;

(3) Tu pourras ouvrir soit la boîte bleue, soit les deux boîtes (la boîte rouge et la boîte bleue), et tu prendras le contenu des boîtes ouvertes ;

(4) Dans la boîte rouge, je mettrai 100 F ;

(5) Avant de remplir la boîte bleue, je ferai une prédiction (que je ne te communiquerai pas) : si je prédis que tu prendras la boîte bleue toute seule, je mettrai 1 000 F dedans ; si je prédis que tu prendras les deux boîtes, je ne mettrai rien dans la boîte bleue.

Que faut-il faire? Ouvrir les deux boîtes, comme les règles nous y autorisent, ou se contenter d'ouvrir uniquement la bleue?

La plupart des gens à qui l'on pose le problème disent qu'il vaut mieux n'ouvrir que la boîte bleue, seule. En effet, cela rapportera 1 000 F – si les prédictions de la fée sont toujours justes –, alors que prendre les deux boîtes – toujours si les prédictions de la fée sont justes – ne rapportera que 100 F.

Mais certaines personnes font un autre raisonnement. Elles disent : «Au moment où je fais

mon choix, les boîtes sont remplies ; donc, à cet instant, je choisis entre prendre le contenu de la boîte bleue uniquement, ou prendre le contenu de la boîte bleue et de la boîte rouge. Je ne veux pas me priver du contenu de la boîte rouge que je sais être de 100 F ! Je dois donc ouvrir les deux boîtes.»

Pour renforcer l'hésitation entre les deux raisonnements, on peut imaginer qu'un arbitre contrôle le jeu : il a connaissance de la prédiction de la fée et il vérifie qu'à l'instant où le joueur choisit, la fée ne change pas le contenu des boîtes. Cela rend difficile le renoncement au contenu de 100 F de la boîte rouge que préconise le premier raisonnement.

Deux raisonnements qui semblent justifiés conduisent à deux conclusions opposées. L'une des hypothèses de départ doit être absurde. C'est sans doute l'hypothèse concernant la prédiction. Donc, le paradoxe de Newcomb montre que la prédiction des choix humains est impossible, même quand celui qui fait les prédictions ne les annonce pas.

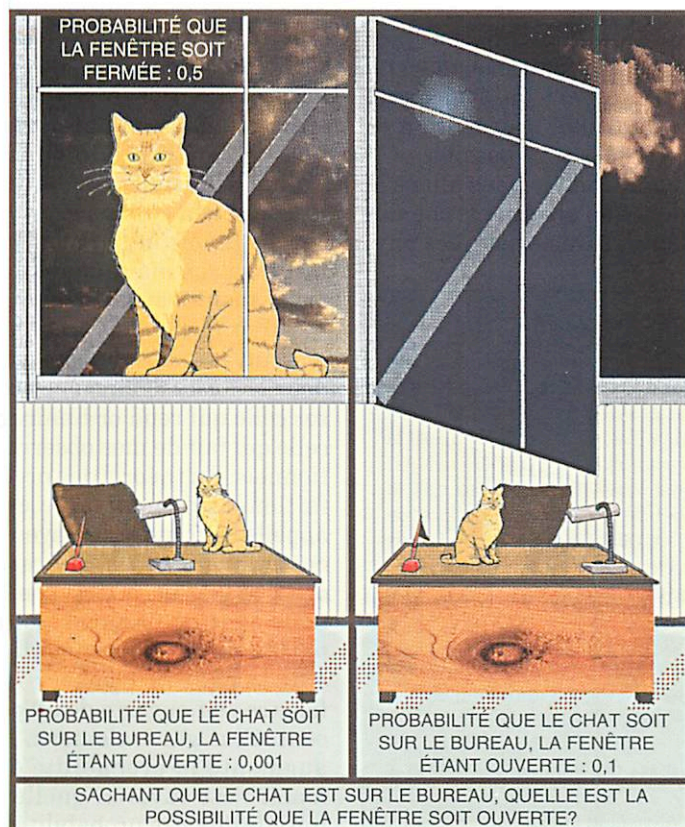
Je trouve cela étrange, car, comme dans le paradoxe de l'apéritif, on pourrait faire intervenir un ordinateur. Supposons en effet que nous ayons programmé un ordinateur pour raisonner comme nous venons de le faire. La conclusion que nous avons obtenue concernant l'imprévisibilité du comportement humain s'appliquerait à notre ordinateur. Nous aurions donc une démonstration que le comportement des ordinateurs est imprévisible, ce qui est incompatible avec la remarque faite plus haut que le comportement d'un ordinateur isolé est toujours prévisible. La contradiction est grave.

Dans le paradoxe de Newcomb avec ordinateur, nous utilisons l'hypothèse qu'il est programmé pour raisonner sur le problème comme nous le faisons nous-mêmes. La contradiction obtenue nous oblige à reconnaître que, cette fois, c'est l'hypothèse d'une programmation possible qui est à rejeter. Nous devons donc conclure du paradoxe de Newcomb avec ordinateur que jamais nous n'arriverons à programmer des machines pour qu'elles raisonnent comme nous. Tiendrions-nous là une preuve de l'absurdité du projet de l'intelligence artificielle qui prétend réaliser des machines ayant des capacités d'intelligence équivalentes à celles des humains ?

Je sais que beaucoup de gens doutent de la faisabilité du projet de l'intelligence artificielle, mais il me semble absurde de croire qu'on peut prouver aussi simplement que cela, et par avance, l'échec de ce programme de recherche, dont les résultats actuels sont loin d'être négligeables. Il y a donc quelque chose qui ne va pas dans ce que je viens de

raconter, et même si je ne sais pas quoi précisément, je ne crois pas que le paradoxe de Newcomb avec ordinateur démontre l'impossibilité de l'intelligence artificielle.

Personne aujourd'hui n'a réussi à formuler clairement la ou les erreurs dans les raisonnements autour du paradoxe de Newcomb, et le vertige dont on est saisi quand on cherche à approfondir le problème n'a d'égal que la conviction très forte que rien ne peut en être tiré concernant le monde réel, ni, en particulier, concernant le projet de l'intelligence artificielle : il y a une inadéquation évidente entre les



2. Je sais qu'une fois sur deux j'oublie de fermer la fenêtre de la cuisine et qu'alors, une fois sur 10, le chat de la voisine vient s'installer sur la table de mon bureau. Lorsque je ferme la fenêtre, le chat n'est presque jamais sur mon bureau, en fait moins d'une fois sur 1 000. Je vois le chat sur mon bureau. Ne dois-je pas en conclure que, très probablement, la fenêtre est ouverte ? La formule de Bayes justifie un tel raisonnement.

moyens utilisés – qui ne portent sur aucun fait réel – et les conclusions qu'on prétend obtenir concernant la liberté humaine ou l'intelligence artificielle. Identifier exactement l'entourloupette n'est pas facile, mais il est certain qu'il y en a une! Les lecteurs intéressés pourront se reporter au livre de W. Poundstone (*Les labyrinthes de la raison*, 1988, éditions Belfond) où le paradoxe de Newcomb est longuement et finement discuté.

Le philosophe que personne ne croit

La troisième difficulté que je souhaite évoquer concernant la prédiction a été récemment avancée par John Leslie, un philosophe spécialisé dans l'étude des problèmes de cosmologie et qui s'est aussi particulièrement intéressé au principe anthropique. Ce principe, qui est l'objet encore aujourd'hui de nombreuses discussions entre astrophysiciens, affirme que : «Lorsque nous réfléchissons à notre situation dans l'Univers, nous ne devons pas nous étonner de nous trouver quelque part où la vie et l'intelligence sont possibles, puisque, si elles ne l'étaient pas, nous ne serions pas là pour le remarquer!» ou, sous forme succincte : «Un observateur doit s'attendre à se trouver là où les observateurs sont possibles.» Il s'agit d'une évidence, mais, comme nous allons le voir plus loin, de l'évidence à l'absurde le chemin est parfois court! Certaines caractéristiques phy-

siques de l'Univers ne s'expliquent bien qu'en utilisant ce principe.

L'argument de l'Apocalypse

Précisons que ce que j'appellerai le paradoxe de Leslie (appelé aussi argument de l'Apocalypse) provoque une réaction violente et immédiate de rejet de pratiquement toute personne à qui on l'explique. Cette réaction est due à la conclusion du raisonnement. Il est étrange que ce soit la conclusion qui gêne, plus que le raisonnement lui-même! Car celui-ci semble être correct dans bien des situations similaires et serait sans doute considéré comme banal s'il ne nous touchait pas de si près, comme nous allons le voir.

Pour introduire le raisonnement de Leslie, considérons d'abord l'histoire du chat de la voisine :

«Je me lève une nuit d'été en me demandant si je n'ai pas oublié de fermer la fenêtre de la cuisine. Je sais que cela m'arrive un jour sur deux. Je sais aussi que, lorsque j'oublie de fermer la fenêtre, le chat de la voisine vient s'installer sur mon bureau dans 10 pour cent des cas. J'évalue aussi que la probabilité pour que le chat de la voisine soit sur mon bureau lorsque la fenêtre est fermée est très faible, disons 0,1 pour cent. J'allume la lumière du bureau, je ne sais pas si la fenêtre de la cuisine est ouverte, mais je vois le chat de la voisine sur mon bureau. N'ai-je pas une bonne raison maintenant de croire que la fenêtre de la cuisine est ouverte?»

Tout le monde s'accorde à croire que oui, car c'est un principe de bon sens que de dire : entre deux théories également probables, je dois préférer celle qui rend ordinaires mes observations, à celle qui fait de mes observations des faits exceptionnels. En théorie des probabilités, cela est parfaitement démontrable et résulte de ce qu'on appelle la formule de Bayes. Un calcul précis avec cette formule indiquerait ici qu'après avoir observé le chat, je dois considérer qu'il y a 99,01 pour cent de chances pour que j'aie laissé la fenêtre de la cuisine ouverte. En résumé: le fait de voir le chat fait passer la probabilité que la fenêtre soit ouverte de 50 pour cent à 99,01 pour cent. Plus généralement, l'observation du chat augmente la probabilité de l'hypothèse que la fenêtre est ouverte, quelle que soit l'évaluation initiale de cette probabilité.

Considérons maintenant deux hypothèses complémentaires que nous appellerons *Théorie A* et *Théorie B* :

– *Théorie A* : l'humanité disparaîtra avant 2150.

La formule de Bayes

Dans notre exemple, il y a deux théories (ou hypothèses) en compétition : la théorie *A* = "La fenêtre est ouverte" et la théorie *B* = "La fenêtre est fermée". La probabilité *a priori* de la théorie *A* est $P_A = 0,5$ et la probabilité *a priori* de *B* est aussi $P_B = 0,5$. On sait par ailleurs que lorsque *A* est vraie, la probabilité que le chat soit sur le bureau est de $Q_A = 0,1$ et que lorsque c'est *B*, elle est de $Q_B = 0,001$. La formule de Bayes indique que la probabilité P'_A que "La fenêtre est ouverte" soit vraie lorsqu'on sait que le chat est sur le bureau (probabilité *a posteriori*) est :

$$P'_A = (P_A Q_A) / (P_A Q_A + P_B Q_B).$$

Bien sûr, la probabilité P'_B que ce soit *B* qui soit vraie lorsqu'on sait que le chat est sur le bureau est :

$$P'_B = P_B Q_B / (P_A Q_A + P_B Q_B).$$

Ici, on obtient :

$$P'_A = 0,5 \times 0,1 / (0,5 \times 0,1 + 0,5 \times 0,0001) = 0,9901 = 99 \text{ pour cent.}$$

Dans le cas de l'Apocalypse, le calcul donne :

$$P'_A = 0,01 \times 0,1 / (0,01 \times 0,1 + (0,99 \times 0,001)) = 0,502512 = 50,25 \text{ pour cent.}$$

Dans le paradoxe des bébés, le calcul donne :

$$P'_A = 0,01 \times 1 / (0,01 \times 1 + 0,99 \times 1/70) = 0,4142 = 41,42 \text{ pour cent.}$$

– *Théorie B* : l'humanité passera le cap de l'année 2150.

Admettons que, dans le cas de l'hypothèse A, un humain sur dix aura connu les années 1990 (ce qui correspond à une estimation raisonnable) et que, dans le cas de l'hypothèse B où les humains continueront à proliférer, un humain sur 1 000 aura connu les années 1990 (la conclusion serait encore plus troublante que celle que nous allons obtenir si on remplaçait le rapport «1 sur 1 000» par «1 sur 100 000» ou par un rapport encore plus petit).

Faisons l'évaluation – optimiste – que la probabilité de la *Théorie A* est de 1 pour cent et que celle de la *Théorie B* est de 99 pour cent. Maintenant posons-nous la question : qu'est-ce qui rend plus probable que je sois ici en 1993? Est-ce la *Théorie A*, pour laquelle «avoir connu les années 1990» est vrai pour une assez grande partie des humains (10 pour cent), ou est-ce la *Théorie B*, pour laquelle «avoir connu les années 1990» est un fait exceptionnel, vrai pour moins de 0,1 pour cent des humains (car beaucoup naîtront après l'an 2000). Comme dans l'histoire du chat de la voisine, entre deux théories je dois préférer la théorie qui fait de mes données des informations ordinaires, à celle qui les fait apparaître comme exceptionnellement rares ; et donc, je dois revoir à la hausse la probabilité de la *Théorie A*. Un calcul précis avec la formule de Bayes donne que

la valeur 1 pour cent en faveur de la *Théorie A* passe à 50,25 pour cent en faveur de la *Théorie A*. La prise en compte du fait que je suis en train de vivre dans les années 1990 fait passer la probabilité d'une Apocalypse proche de 1 pour cent à 50,25 pour cent!

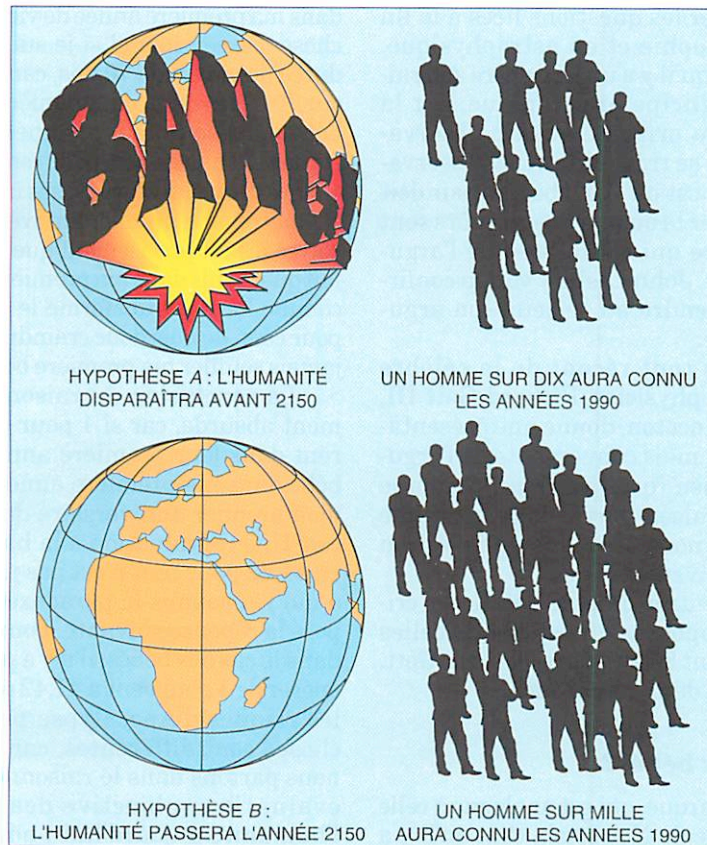
Très peu de gens admettent que l'argument de l'Apocalypse est juste : comment le seul fait de tenir compte que je suis vivant en 1993 pourrait-il justifier le passage d'une évaluation de 1 pour cent pour la *Théorie A* à une évaluation de plus de 50 pour cent, qui signifierait que la fin de l'humanité est imminente?

John Leslie, qui fut le premier à publier ce raisonnement en 1989, indique qu'il a été proposé par Brandon Carter – l'astrophysicien inventeur de l'expression «principe anthropique» – dans

une conférence faite en 1983, mais que ce dernier par peur de choquer – ou parce qu'il n'arrivait pas à être complètement certain de la justesse de son raisonnement? – n'avait pas repris le détail du raisonnement dans le texte écrit de sa conférence.

L'argument apparaît absurde à la plupart des gens à qui vous l'exposez pour la première fois. Il les conduit même à douter de votre bonne santé mentale si vous insistez en disant que vous pensez qu'il s'agit d'un argument à prendre au sérieux.

Le lecteur choqué du raisonnement et qui pense disposer de bons arguments pour le contrer doit se méfier. Il



3. Quelle que soit l'évaluation *a priori* que nous fassions de la probabilité de l'hypothèse A : «l'Apocalypse se produira avant 2150», l'utilisation de la formule de Bayes – comme dans le cas du chat sur la table – fait croître cette évaluation quand nous prenons en compte l'information que nous sommes en 1993. Par exemple, en supposant – ce qui est optimiste – que la probabilité de A est initialement de un pour cent, nous arrivons à la réévaluation de $(1/100 \times 1/10) / [1/100 \times 1/10 + 99/100 \times 1/1\ 000] = 0,502512 = 50,25$ pour cent. La probabilité de A, lorsque nous tenons compte de notre position dans le temps, passe donc de un pour cent à 50,25 pour cent.

est très peu probable qu'il ait trouvé un contre argument original que John Leslie n'ait pas déjà décortiqué, car, comme je le disais plus haut, l'article détaillé du philosophe publié dans la prestigieuse revue *Mind* (vol. 101, n° 403, juillet 1992, pp. 521-540) donne des réponses assez bonnes – à mon sens – à toutes les critiques élémentaires. Si, malgré tout, des lecteurs pensent disposer d'arguments clairs contre le raisonnement de Leslie, je leur suggère d'entrer en contact directement avec lui pour les lui expliquer. Son adresse, qu'il m'a autorisé à publier, est : Département de philosophie, Université de Guelph, Guelph, Ontario, N1G2W1, Canada. Il parle le français et se fera un plaisir de répondre comme il a répondu aux longues lettres que je lui ai envoyées. John Leslie prépare aussi un livre où il traitera plus généralement de toutes les façons rationnelles d'aborder les questions liées à la fin du monde en philosophie et en astrophysique.

John Leslie note qu'il y a un lien entre son raisonnement et le principe anthropique, car la forme généralisée du principe qu'«Un observateur doit s'attendre à se trouver là où les observateurs sont possibles» est qu'«Un observateur doit s'attendre à se trouver là où les observateurs sont le plus probables», ce qui est la base de l'argument de l'Apocalypse. John Leslie y voit la confirmation qu'il faut prendre au sérieux son argument de l'Apocalypse.

Dans un article tout récent de la célèbre revue *Nature*, l'astrophysicien Richard Gott III, de l'Université de Princeton, donne une présentation assez différente, mais convergente de l'argument de l'Apocalypse (qu'il relie au principe copernicien : «sans raison particulière, c'est une erreur de croire que nous occupons une position privilégiée dans l'Univers»).

Je ne veux pas entrer dans le détail des critiques qui ont été opposées à Leslie et de celles – prévisibles – qui vont l'être à l'article de R. Gott. Je me contenterai de deux remarques.

Qu'en pensent les bébés?

La première remarque, qui est analogue à celle faite pour le paradoxe de Newcomb, c'est qu'il y a une disproportion grave entre les moyens mis en œuvre dans le raisonnement – moyens dérisoires n'invoquant aucun fait matériel nouveau – et la conclusion obtenue qui, elle, concerne notre avenir proche et qui, concrètement, signifie que les risques nucléaires, les risques dus à la pollution ou aux épidémies doivent être pris plus au sérieux qu'ils ne le sont lorsqu'on ne fait pas le raisonnement.

Le deuxième point qui me fait espérer que le raisonnement de Leslie est inacceptable est ce que j'appelle le paradoxe des bébés.

Imaginons que les données suivantes sont exactes (elles ne le sont pas précisément pour la France, mais sont assez proches des données réelles).

- 1 pour cent des bébés meurent dans leur première année de vie.

- Pour les 99 pour cent qui passent leur premier anniversaire, la durée de vie moyenne est de 70 ans.

En copiant Leslie et en imaginant que je suis un bébé de moins de 1 an, je peux alors faire le raisonnement suivant :

«Le fait que je sois dans ma première année de vie est ordinaire si je suis un humain qui n'atteint pas son premier anniversaire ; en revanche, être dans ma première année de vie est pour moi quelque chose d'exceptionnel si je suis un humain dont la durée de vie est de 70 ans, car je n'ai qu'une chance sur 70 d'être dans ma première année. Si je suis un bébé dans sa première année de vie, je dois donc revoir à la hausse le 1 pour cent de chances (de malchances!) que les statistiques m'attribuent de ne pas connaître mon premier anniversaire. La formule de Bayes me dit précisément que prendre en compte le fait que je suis dans ma première année, en calculant comme Leslie, transforme le 1 pour cent en 41,42 pour cent. Je dois donc craindre sérieusement de ne jamais souffler ma première bougie.»

La conclusion du raisonnement est résolument absurde, car si 1 pour cent des bébés meurent dans leur première année et que je suis un bébé dans ma première année de vie, j'atteindrai mon premier anniversaire dans 99 pour cent des cas. Une réévaluation à la hausse du 1 pour cent en 41,42 pour cent n'est pas justifiée. John Leslie, à qui j'ai soumis le paradoxe des bébés, m'a proposé la réponse suivante. Pour lui, il est exact que, dans le cas des bébés, il n'y a pas de réévaluation à opérer de 1 pour cent à 41,42 pour cent. Mais, pour les risques d'Apocalypse prochaine, dit-il, les choses sont différentes, car la probabilité dont nous partons dans le raisonnement résulte d'une évaluation subjective des risques que court l'humanité aujourd'hui ; nous sommes donc en droit de modifier cette évaluation approximative qui, contrairement au cas des bébés, n'est pas le produit d'études statistiques objectives.

Je laisse les lecteurs évaluer si cette réponse est satisfaisante, et je m'excuse auprès d'eux d'avoir peut-être jeté le trouble dans leur esprit en exposant des paradoxes dont je connais le pouvoir obsessionnel.