Cahiers de Psychologie Politique

N° 33 | 2018
Quantification et quantité Juillet 2018

Rationalité et représentation numérique

François LEPAGE

Édition électronique :
URL : https://cpp.numerev.com/articles/revue-33/1925-rationalite-et-representation-numerique
DOI : numerev_1175
ISSN : 1776-274X
Date de publication : 16/07/2018

Cette publication est sous licence **CC BY-NC-ND** (Attribution - No commercial - No derivatives).

Pour **citer cette publication** : LEPAGE, F. (2018) Rationalité et représentation numérique. *Cahiers de Psychologie Politique*, (33). https://doi.org/10.34745/numerev_1175

DOSSIER : QUANTIFICATION ET QUANTITE

« On peut faire dire n'importe quoi aux statistiques. » Combien de fois n'avons-nous pas entendu, sous une forme ou une autre, cette affirmation péremptoire qui exprime une méfiance profonde envers la mathématisation, voire la simple formalisation, des lois des sciences humaines. Il est difficile de prêter un sens univoque à une telle formule mais je pointerais vers une explication du genre suivant. À partir du moment où on attache des valeurs numériques à des variables en sciences humaines, on glisse vers une forme de positivisme et nos théories deviennent des caricatures naïves des phénomènes complexes dont elles sont censées rendre compte.

Mots-clés :

DOSSIER : QUANTIFICATION ET QUANTITE

François Lepage est professeur honoraire, spécialiste de logique et épistémologie au département de philosophie de l'université de Montréal. Il est l'auteur d'Éléments de logique contemporaine aux Presses Universitaires de Montréal (2011) et de Logique partielle et savoir publié chez Bellarmin Vrin (2006).

SOMMAIRE

1. Motivation

2. Le concept de préférence et sa formalisation

3. Le concept d'utilité attendue et sa formalisation

4. Questions, problèmes et remarques

5. Contre-exemples

5.1 Le paradoxe d'Allais

5.2 Violation de la transitivité des préférences

5.3 La complétude est-elle crédible ?

6. Faut-il mettre les théories normatives à la poubelle ?

1. Motivation

« On peut faire dire n'importe quoi aux statistiques. » Combien de fois n'avons-nous pas entendu, sous une forme ou une autre, cette affirmation péremptoire qui exprime une méfiance profonde envers la mathématisation, voire la simple formalisation, des lois des sciences humaines. Il est difficile de prêter un sens univoque à une telle formule mais je pointerais vers une explication du genre suivant. À partir du moment où on attache des valeurs numériques à des variables en sciences humaines, on glisse vers une forme de positivisme et nos théories deviennent des caricatures naïves des phénomènes complexes dont elles sont censées rendre compte.

À l'extrême opposé, se trouve un postulat épistémologique radical : les régularités empiriques observables en sciences humaines doivent pouvoir être représentées par des lois formelles et, si possible, numériques sous peine soit de perdre leur caractère scientifique et d'être carrément incompréhensibles, soit de retomber dans la psychologie populaire illustrée par ces nombreux ouvrages de croissance personnelle.

Dans les pages qui suivent, je propose d'étudier un cas simple de formalisation et de représentation numérique, celui des échelles de préférence et de choix rationnel et d'en faire un examen critique.

Je ne suis ni un spécialiste des sciences humaines ni un spécialiste des mathématiques de la prise de décision. On ne trouvera dans cette intervention aucune contribution originale. Je travaille en logique philosophique et j'ose seulement penser qu'un regard ni de trop près ni de trop loin peut apporter matière à réflexion. Ce texte comporte deux parties. La première est une présentation semi formelle de l'axiomatisation des notions de préférences et de l'utilité attendue des actes d'un agent supposé être rationnel. La seconde est une analyse critique de cette axiomatisation et de ses limites.

2. Le concept de préférence et sa formalisation

Un des concepts clés que nous utilisons dans nos tentatives de description du comportement rationnel est celui de *préférence*. On peut considérer ce concept comme primitif tout comme on peut le considérer comme une abstraction dérivant de l'observation empirique du comportement des agents. Ma chatte gratte à la porte menant au jardin. Je lui ouvre et je la vois hésiter. Il pleut et la température est de deux degrés. Elle fait volte-face et retourne se coucher sur son fauteuil habituel. Elle préfère l'option de rester à l'intérieur à celle de sortir. Notre compréhension des comportements des agents contrairement à ceux des non-agents fait appel à ce concept. Nous ne dirions pas que l'allumette que j'ai frottée préfère s'enflammer ou encore que le caillou

qui m'échappe des mains préfère tomber. Le concept de préférence est une relation entre options qui s'offrent à des agents « rationnels ».

Il y a plusieurs façons plus ou moins équivalentes de formaliser la relation de préférence. Nous commencerons avec l'une des plus simples qui permet quand même de présenter des problèmes philosophiquement pertinents. La notion d'agent est primitive : un agent peut être un animal, un être humain, une collectivité, un ordinateur, etc. On se donne un ensemble d'options. Le statut ontologique de ces options ne fait pas l'unanimité et, pour certains, ce sont des lots, quelquefois des montants d'argent. La majorité des exemples que l'on retrouve dans les manuels de théorie de la décision, de la théorie des jeux, etc., sont des exemples monétaires. On ne peut s'empêcher de penser qu'il y a là une source de la méfiance envers les échelles de préférence quantifiées : cette approche monétaire connote l'idée que tout se vend et tout s'achète…

Pour de nombreux philosophes, il est plus pratique de considérer les options comme des propositions, en fait, les propositions affirmant que l'agent obtient tel ou tel lot, et c'est ce que nous ferons ici. Nous nous autoriserons quand même, par souci de simplicité, à parler de lots. Pour les besoins de ce texte, nous supposerons que les propositions de base sont en nombre fini. À chaque agent est associée une relation entre les lots : la relation de préférence faible : *A*â¿*B* qui se lit l'agent préfère faiblement *A* à *B*. Le mot « faible » est utilisé parce que nous voulons pouvoir définir l'indifférence : *A*ï¾*B* qui se lit l'agent est indifférent entre *A* et *B* et se définit par *A*â¿*B* et *B*â¿*A*. On a, bien sûr, *A*ï¾*A* et la relation d'indifférence tout comme celle de préférence faible sont réflexives.

La plupart des auteurs, comme Luce et Raiffa (1985), prennent le concept de préférence faible comme primitif.

Nous allons maintenant nous intéresser à une propriété sensiblement plus problématique, la *transitivité* des préférences faibles :

A â¿ *B* et (*B* â¿ *C*), alors (*A* â¿ *C*).

Par exemple, vous préférez « utiliser votre vélo pour aller au travail » à « prendre le métro pour aller au travail » et vous préférez « prendre le métro pour aller au travail » à « prendre votre auto pour aller au travail ». Donc, vous préférez « utiliser votre vélo pour aller au travail » à « prendre votre auto pour aller au travail ».

Cette propriété, nous le verrons plus loin, est empiriquement contestable.

Enfin, une troisième propriété que nous retiendrons est la complétude : pour tout *A*, *B*, *A*â¿*B*ou *B*â¿*A*, qui est également problématique[1]. En effet, cette propriété présuppose que toutes les options possibles sont comparables.

En théorie des ensembles « â¿ », telle qu'axiomatisée ici, est un préordre total. L'intuition naïve associée à la relation de préférence faible et à celle d'indifférence est celle d'une valeur *subjective* plus ou moins grande que l'agent attache aux options.

C’est presque un pléonasme d’affirmer que si l’agent préfère *A* à *B*, c’est qu’il accorde une *plus grande* valeur à *A* qu’à *B*. Il est tentant de représenter la valeur de chacune des options par un nombre : à chaque option A_i est attaché un nombre réel $U(A_i)$ qu’on appelle l’utilité de cette option. On a

$A \hat{a}\square\!\!\!

L'ensembledes actes est aussi un ensemble de propositions, l'ensemble des propositions que l'agent croit pouvoir rendre vraies. Par exemple, l'agent prend son parapluie, l'agent se rend à la plage, l'agent achète des actions de Monsanto, l'agent claque des doigts pour éteindre le soleil. À chaque acte *X* est associée une conséquence qui dépend de l'état du monde. La connaissance de l'état du monde est cependant habituellement incertaine comme dans l'exemple de Londres ci-dessus : l'agent se demande s'il est dans un monde où il va pleuvoir ou bien s'il est dans un monde où il ne pleuvra pas. Prendre son parapluie dans un monde où il pleuvra n'a pas la même conséquence que prendre son parapluie dans un monde où il ne pleuvra pas, d'où la nécessité d'utiliser des *probabilités subjectives*.

On définit alors l'utilité attendue d'un acte *X* de la manière suivante :

$UA(X) = \Pr_x(A_1)U(A_1) + \ldots + \Pr_x(A_n)U(A_n)$

où $\Pr_x(A_i)$ est la probabilité de l'option A_i si l'agent accomplit l'acte *X*. Cette formule est la traduction numérique de l'intuition pascalienne : l'utilité attendue d'un acte potentiel *X* est la somme pondérée des utilités des conséquences, les coefficients de pondération étant les probabilités que chaque conséquence se produise si l'acte est accompli (la somme des probabilités est, bien sûr, de 1. Un exemple extrêmement banal est celui de l'achat d'un billet de loto. Supposons que les seules utilités soient les montants d'argent (nous verrons plus loin que cette hypothèse est presque toujours simplificatrice). L'utilité attendue de l'achat d'un billet de loto est

$UA(X) = \Pr_x$(le billet est gagnant) x (le montant gagnant – le prix du billet) + $\Pr_x$(le billet est perdant) x (– le prix du billet)

Cette notion d'utilité d'un acte souffre d'un défaut : les valeurs $U(A_i)$ n'ont, en général, qu'une signification ordinale. Dans l'exemple du billet de loto, on suppose que les valeurs ordinales et cardinales coïncident, un million d'euros valant un million de fois la valeur d'un euro.

Prenons, encore une fois, l'exemple du parapluie et supposons que les préférences de l'agent sont représentées par la matrice suivante :

	p	non *p*
X	V_1	V_2
non X	V_3	V_4

Avec V_1 l'agent prend son parapluie et il pleut, V_2 l'agent prend son parapluie et il ne pleut pas, etc. Supposons que l'ordre des préférences de l'agent soit V_1 â¿ V_4 â¿ V_2â¿V_3 et que la probabilité qu'il pleuve soit de 0,4 et donc que celle qu'il ne pleuve pas soit de 0,6. Les deux matrices d'utilité suivante respectent l'ordre des préférences de l'agent :

	p	non *p*
X	5	-5

non X	-6	4

	p	non p
X	5	-2
non X	-4	2

La première nous donne les utilités attendues

$UA_1(X) = (5x0,4)+(–2x0,6) = 0,8$ et $UA_1(\text{non } X) = (–4x0,4) + (2x0,6) = 0,4$

alors que la seconde nous donne

$UA_2(X) = (5x0,4)+(–5x0,6) = –1$ et $UA_2(\text{non X}) = (–6x0,4)+(4x0,6) = 0$.

Bien que les deux matrices d'utilité soient compatibles avec les préférences de l'agent, les utilités attendues sont incompatibles : on a $UA_1(X) > UA_1(\text{non } X)$ et $UA_1(X) < UA_1(\text{non } X)$. Nous devons donc conclure que la donnée d'une fonction d'utilité définie à partir d'une relation de préférence plus la donnée d'une fonction de probabilité sur les états du monde ne permettent pas, en général, de définir l'utilité attendue d'un acte.

La question est de savoir s'il est possible de définir une notion d'utilité attendue qui ait une valeur cardinale fiable. La réponse est oui et l'une des façons d'y parvenir est d'introduire la notion de *loterie* comme l'ont fait von Neumann et Morgensten (1944).

Une loterie est simplement une liste de probabilités $L = [p_1;p_2;...;p_n]$où chaque p_i est la probabilité de l'option A_i. La relation de préférence primitive associée à l'agent n'est plus entre les options mais entre les loteries.

Tout comme nous avons eu besoin d'axiomes pour caractériser les propriétés des préférences entre les options, nous avons besoin d'axiomes pour caractériser les préférences entre les loteries.

(1') *Complétude* : Si L et L' sont deux loteries, alors Lâ¿L' ou L' â¿L.

D'un point de vue méthodologique, les agents ont des préférences sur *toutes* les loteries et sont donc en mesure d'exprimer leur préférence entre n'importe quelle paire.

(2') *Transitivité* : Si Lâ¿L' et L' â¿L'' alors L' â¿L''.

Les deux autres axiomes sont un peu plus techniques. On suppose d'abord que les loteries sont elles-mêmes des options et qu'elles ont une probabilité. Cela permet de définir les loteries composées.

Prenons le cas simple de deux loteries, $L = [p_1;(1- p_1)]$ et , L ayant la probabilité x et L' ayant la probabilité $(1 - x)$. La loterie composée sera notée $xL + (1-x)L' =$. Nous arrivons ainsi au troisième axiome :

(3') *Continuité* : Soit trois loteries telles que $L \succeq L' \succeq L''$. Il existe alors un nombre réel $x \in [0,1]$ tel que $L' \sim (xL + (1-x) L'')$.

Voici un exemple très simple.

Un agent reçoit un héritage de trois cent mille euros. Il réfléchit à l'usage qu'il pourrait faire de cette somme. Il envisage alors trois possibilités. La première est l'achat d'un (modeste) appartement au centre ville, ce qui le rapprocherait de son lieu de travail. Fini les mornes trains de banlieue. La seconde est l'achat d'une maison de campagne, vieux rêve qui deviendrait réalité. Enfin la troisième, est tout simplement d'investir dans un fonds de placement à faible risque dont les intérêts annuels suffiraient à payer son loyer. Remarquons que si la valeur monétaire de chacun des lots est la même, trois cent mille euros, l'utilité de chacun de ces lots peut varier selon la subjectivité de l'agent. Supposons que, après mûre réflexion, l'agent préfère l'option *A* « achat d'un appartement » à l'option *M* « achat d'une maison de campagne « elle même préférée à l'option *P* « fonds de placement ».

Soit L_A, L_M et $L_P$les loteries élémentaires qui donnent respectivement avec certitude chacun des lots. Nous avons donc $L_A \succeq L_M \succeq L_P$. L'axiome de continuité implique qu'il existe une probabilité x telle que $L_P \sim (xL_V + (1-x)\ L_C)$. Autrement dit, l'agent est indifférent entre l'achat d'une maison de campagne et un ticket de loterie qui lui donnera un appartement avec la probabilité x ou un fonds de placement avec la probabilité $(1-x)$.

Enfin, le quatrième axiome s'énonce ainsi si une loterie est préférée à une autre, aucune composition avec une troisième ne renversera cette préférence.

(4') Indépendance : pour tout x différent de 0 et plus petit ou égal à 1 et pour toutes loteries L,L' et L'', $L \succeq L'$ si, et seulement si, $(xL + (1-x)L'') \succeq (xL'+(1-x)L'')$.

Nous arrivons ainsi au terme de notre présentation formelle :

Théorème de représentation de l'utilité attendue (von Neumann et Morgenstern).

Si les préférences sur les loteries d'un agent satisfont (1')-(4'), alors il est possible de définir une fonction d'utilité UA[2]telle que

$L \succeq L'$ si, et seulement si, $UA(L) \geq UA(L')$[3].

4. Questions, problèmes et remarques

On peut schématiser la démarche exposée ci-dessus comme comportant trois phases.

(a) La première phase est l'identification des notions préthéoriques que l'on prête aux agents « rationnels » : préférences sur des « lots », probabilités subjectives ou croyances attribuées par les agents aux états possibles du monde, possibilité d'actions volontaires visant à obtenir le meilleur lot, etc.

(b) La seconde phase est la formalisation par abstraction et l'axiomatisation de ces différentes notions. L'exemple le plus simple est celui de préférence. La notion préthéorique de préférence suggère d'elle-même l'existence d'une certaine *relation d'ordre* sur un ensemble de lots. Ici, nous avons une relation d'ordre très simple, celle qu'on appelle un *préordre total*.

(c) La troisième phase est celle de la démonstration de l'existence d'un *théorème de représentation*, c'est-à-dire la démonstration de l'existence d'une fonction numérique qui assigne une valeur à chaque lot de sorte que l'ordre de préférence entre les loteries est le même que l'ordre numérique entre les utilités attendues de ces mêmes loteries : la délibération rationnelle d'accomplir tel ou tel acte est ramenée à une simple comparaison entre nombres réels calculés par l'utilité attendue des loteries primitives, elles-mêmes entièrement déterminées par l'ordre ordinal des préférences sur les loteries.

Une première remarque s'impose. Le passage de (b) à (c) semble immunisé contre toute critique. La construction ne fait intervenir que des mathématiques banales au dessus de tout soupçon. S'il n'est pas possible d'établir un théorème de représentation, comme ce fut le cas plus haut en axiomatisant la relation de préférence uniquement sur les lots, c'est que l'axiomatisation n'est pas assez développée ou n'est tout simplement pas adéquate.

On ne saurait en dire autant du passage de (a) à (b). Par exemple, que la relation d'indifférence soit réflexive est difficilement contestable, mais qu'en est-il de l'hypothèse que la préférence faible soit une relation totale ? Il semble probable que ce soit un choix méthodologique. Enfin, concernant la transitivité, on a des exemples empiriques qui montrent que certains agents ne la respectent pas.

Ceci nous amène à la question du statut des théories axiomatisées de l'utilité attendue. Considéré comme une « explication » à la Carnap, c'est-à-dire comme une substitution de concepts purement formels explicites à des notions vagues, le point de vue ne peut être que *normatif*. Cette situation est similaire à celle du statut de la notion d'ensemble. On connaît le sort qu'a subi la théorie naïve des ensembles après la découverte du fameux paradoxe de Russell. On croyait que l'on pouvait « rassembler » n'importe quoi : une pomme, la lune, le concept d'étoile, etc. Aujourd'hui, si on pose la question de la nature des ensembles, on se verra probablement répondre que ce sont les objets de

base de n'importe quelle structure mathématique qui est un modèle de la théorie axiomatique de Zermelo-Frankel (par exemple). La théorie axiomatisée de l'utilité attendue dans le cadre de l'explication de l'action rationnelle est une explication de ce que c'est que de prendre une décision rationnelle. Si le processus qu'utilise un agent pour prendre une décision est en violation plus ou moins flagrante avec les axiomes, ce n'est pas une décision rationnelle, point final.

Si un agent utilise une théorie naïve des ensembles usant du principe que toute propriété détermine un ensemble, il risque de rencontrer une contradiction.

Si un agent utilise une théorie de la décision qui ne respecte pas les axiomes de l'utilité attendue, il risque d'avoir de mauvaises surprises (nous verrons plus loin des exemples de mauvaises surprises).

À cette approche, s'opposent les théories *descriptives* : la théorie de la décision rationnelle doit décrire comment les agents prennent réellement leurs décisions, ce à quoi nous ferons allusion plus loin.

5. Contre-exemples

5.1 Le paradoxe d'Allais

Historiquement, le plus bel exemple de comportement irrationnel (au sens de la théorie normative), est le paradoxe d'Allais. Il s'agit d'un exemple concret où certains agents violent plus ou moins consciemment l'axiome d'indépendance.

On présente aux agents successivement deux paires de loteries, L_1,L_2 et L_3,L_4, avec les lots A_1 = 500 millions d'euros, A_2 = 100 millions d'euros et A_3 = 0 euro.

La première paire est :

L_1 = [0;1;0].

L_2 = [0,1;0,89;0,01].

La seconde paire est :

L_3 = [0;0,11;0,89].

L_4 = [0,1;0;0,9].

L_1 est une loterie qui donne de façon certaine cent millions d'euros à l'agent puisque la probabilité d'obtenir cinq cent millions tout comme de ne rien obtenir est nulle.

L_2 est une loterie qui donne cinq cents millions avec une probabilité de 0,1, cent millions

d’euros avec une probabilité de 0,89 rien du tout avec une probabilité nulle.

L_3 est une loterie qui donne cinq cents millions avec une probabilité nulle, cent millions d’euros avec une probabilité de 0,11 rien du tout avec une probabilité 0,89.

L_3 est une loterie qui donne cinq cents millions avec une probabilité de 0,1, cent millions d’euros avec une probabilité nulle et rien du tout avec une probabilité 0,9.

Les expériences empiriques montrent que de nombreux agents préfèrent L_1 à L_2 et L_4 à L_3. Les explications qu’ils fournissent suivent généralement les lignes suivantes. L_1 est préférée à L_2 car avoir 100 millions d’euros de façon certaine est préférable à une loterie qui peut rapporter 500 millions d’euros mais avec laquelle on a une chance sur cent de ne rien avoir. L_4 est préférée à L_3 parce que les probabilités d’avoir zéro euro sont presque les mêmes et donc les probabilités de gagner sont presque les mêmes mais dans L_4 le montant est cinq fois plus élevé. Cette explication est relativement naturelle mais est incompatible avec la théorie de l’utilité attendue.

Remarquons d’abord que cette incompatibilité n’a rien à voir avec les valeurs des gains. Soit u_1,u_2 et u_3 les utilités respectives d’obtenir 500 millions d’euros, 100 millions d’euros et zéro euro. On ne fait aucune hypothèse sur ces utilités : u_1,u_2 et $u_3$peuvent prendre n’importe quelle valeur numérique. Les utilités attendues des quatre loteries sont les suivantes :

$UA(L_1) = u_2$

$UA(L_2) = 0{,}1u_1 + 0{,}89u_2 + 0{,}01u_3$

$UA(L_3) = 0{,}11u_2 + 0{,}89u_3$

$UA(L_4) = 0{,}1u_1 + 0{,}9u_3$

Un calcul élémentaire nous donne :

$UA(L_1) > UA(L_2)$ si, et seulement si, $u_2 > 0{,}1u_1 + 0{,}89u_2 + 0{,}01u_3$ si, et seulement si, $0{,}11u_2 > 0{,}1u_1 + 0{,}01u_3$

et

$UA(L_4) > UA(L_3)$ si, et seulement si$UA(L_4) = 0{,}1u_1 + 0{,}9u_3 > 0{,}11u_2 + 0{,}89u_3$ si, et seulement si $0{,}1u_1 + 0{,}01u_3 > 0{,}11u_2$

ce qui est une contradiction flagrante quelles que soient les valeurs de u_1,u_2 et u_3. Cette contradiction ne dépend donc aucunement de l’ordre entre les utilités u_1,u_2 et u_3 : la contradiction provient d’une violation de l’axiome d’indépendance.

Bref, quelle que soit l’attitude d’un agent vis à vis de l’argent, s’il viole

L_1â¿L_2 si, et seulement si, L_3â¿L_4

il ne maximise pas une fonction d'utilité telle que définie plus haut.

Le problème provient du fait que la théorie de l'utilité attendue ne prend pas en compte la sensibilité des agents envers le risque. Le paradoxe d'Allais est un cas particulier où une des loteries rapporte un lot de façon certaine ce qui, pour de nombreux agents, constitue un « plus » indépendant de l'utilité attendue.

5.2 Violation de la transitivité des préférences

Voici un exemple empirique de violation de la transitivité des préférences dû à Tversky (1969). L'expérience fut menée à l'Université de Harvard sur un groupe d'étudiants à propos des préférences associées ente deux loteries. On considère les cinq loteries suivantes :

Les probabilités n'étaient pas données sous forme numérique mais sous forme de portion de « tarte » pour éviter que les étudiants fassent un calcul. On demandait aux étudiants de donner leurs préférences entre deux loteries.

Les résultats on montré que, pour des loteries adjacentes, la préférence allait à celle qui avait le gain potentiel le plus élevé, les probabilités étant perçues comme passablement équivalentes. On avait donc généralement L_iâ» L_{i+1}. Lorsque les probabilités étaient passablement différentes, le gain le plus probable était le facteur dominant du choix. On se retrouve donc avec une structure de préférence telle que L_1â» L_2,...,L_4â» L_5,L_5â» L_1. Cette circularité est en violation flagrante avec la transitivité des préférences.

5.3 La complétude est-elle crédible ?

Y a-t-il des lots tellement détestés qu'un agent ne peut les comparer ? Supposons que vous êtes très allergique au poisson, que sa consommation peut vous être mortelle. Vous êtes au restaurant et, au menu, il y a du poulet, du bœuf, du lieu et du grondin. Préférez-vous le lieu au grondin, le grondin au lieu ou êtes-vous indifférent ? Vous ne préférez ni le lieu au grondin ni le grondin au lieu. Doit-on conclure que vous êtes indifférent ? Supposons que oui. Supposons également que vous êtes un grand amateur de bourgogne blanc. Dans ce cas, vous préférez le grondin avec un Pouilly-Fuissé au lieu avec de l'eau du robinet. Cela semble tout à fait invraisemblable.

Un agent rationnel est-il prêt à acheter une loterie dans laquelle un des lots est sa propre mort ou celle d'un être cher ? La réponse qui nous vient à l'esprit est non. Pourtant, nous sommes presque tous prêts, dans certaines circonstances, à acheter pour 900 euros un ticket de loterie [p;(1–p)] avec A_1 = un aller-retour Paris-New York et

A_2 = l'avion se crashe avec p = 0,9999917.Cette probabilité est farfelue, je l'ai trouvée en tapant « probabilité qu'un avion s'écrase » sur un moteur de recherche ! Est-ce le nombre de crashs par vol ? Par heures de vol ? Y a-t-il une pondération par compagnie ? Tout ce dont nous avons besoin pour la crédibilité de cet exemple est que la probabilité subjective d'un crash mortel soit très petite par rapport à celle de faire une traversée de l'Atlantique, certes désagréable, mais sécuritaire. La morale de l'histoire demeure la même quelle que soit la probabilité qu'un avion s'écrase, pourvu qu'elle soit faible. Tous ceux qui prennent l'avion achètent un ticket semblable.

6. Faut-il mettre les théories normatives à la poubelle ?

Les théories normatives échouent à rendre compte des comportements par ailleurs perçus comme rationnels. N'est-ce pas ce que font les gens qui mettent leur pécule à la Caisse d'épargne plutôt que de se doter d'un portefeuille d'actions qui, sur trente ans, rapportera *presque* certainement au moins 10 fois plus. Mais lorsque le Krach arrivera, ils auront le sourire aux lèvres.

La théorie normative qu'est celle de l'utilité attendue a au moins une fonction sur laquelle tout le monde s'accordera, celle de montrer ses limites. Si les agents ne maximisent pas une fonction d'utilité attendue et que leur comportement est quand même perçu comme rationnel - ces agents ne font pas n'importe quoi -, qu'est-ce qu'ils font ? Ils font ce qu'ils croient être dans leur meilleur intérêt[4], ce qu'ils croient être la *meilleure chose* à faire. Bref, ils maximisent une certaine fonction qui n'est pas une fonction d'utilité telle que définie ci-dessus. Entre maximiser une fonction et agir de façon aléatoire, il ne semble pas y avoir beaucoup d'espace. Et puis, sans la théorie de l'utilité attendue, il n'y aurait jamais eu le paradoxe d'Allais. Sans le paradoxe d'Allais, les travaux de Kahneman et Tversky (1979) sur la théorie des prospects n'auraient jamais vu le jour.

Revenons maintenant sur la transitivité. Dans l'exemple ci-dessus, l'hypothèse est que les montants d'argent sont exactement corrélés à leur utilité : $U(x\$) \geq U(y\$)$ ssi $x \geq y$. Le bris de transitivité ne peut provenir que d'un problème de probabilité. Remarquons que ce n'est pas sur l'incapacité, pour l'agent, d'évaluer avec précision la valeur numérique que représente la portion de tarte, que repose le problème.

Supposons que nous ayons la suite suivante : $p_1u_1 > p_2u_2$, $p_2u_2 > p_3u_3$ et $p_3u_3 > p_1u_1$. La fonction de croyance d'un agent qui, comme les étudiants de Harvard, adhère à cette suite de choix ne peut pas être une fonction de probabilité. Et si la fonction de croyance d'un agent n'est pas une fonction de probabilité, on peut lui monter une loterie qu'il trouvera acceptable à un certain coût, mais qui lui sera défavorable quel que soit l'état du monde. (Voir Skyrms (1986)). C'est ce qu'on appelle un *dutchbook*. Pour pouvoir monter un dutchbook, il faut cependant que le comportement de l'agent obéisse aux axiomes gouvernant les préférences et l'utilité attendue et nous avons vu que c'est,

pour le moins, une hypothèse très contestable et, dans de nombreuses situations réelles et non du genre « expérience de pensée », falsifiée.

Que conclure ? Les concepts à l'œuvre dans la description des activités des agents « rationnels » sont d'une grande complexité. L'adoption de positions intransigeantes et dogmatiques n'est pas une façon de faire avancer l'analyse philosophique de cette question centrale. Oui, les théories normatives ont de sévères limites. Les comportements des agents les falsifient régulièrement. Elles nous apprennent cependant beaucoup de choses. Les rejeter du revers de la main ne nous fait pas avancer dans la compréhension du comportement humain. Soyons pragmatiques, voire opportunistes. Tâtons de toutes les approches. Comme m'a dit un ami haïtien, « Catholique le matin, évangéliste l'après-midi et, le soir, tous au vaudou ». Un être rationnel ne prend pas de risque !

1 Nous ne retenons que ces trois propriétés très simples mais suffisantes pour illustrer notre analyse. Pour une présentation de différentes axiomatisations des préférences, voir Bouyssou *et. al.*

2 Le théorème de représentation ne détermine pas une utilité attendue unique mais une classe de fonctions obtenues les unes à partir des autres par une transformation linéaire comme on obtient les degrés Fahrenheit à partir des degrés Celsius.

3 Pour une approche simple de la démonstration purement mathématique, voir Shih En Lu dans la bibliographie.

4 Il faut être prudent avec la notion d'intérêt qui a presque toujours une connotation égoïste. Le meilleur intérêt d'un agent n'est pas toujours, presque jamais, égoïste ou pire, monétaire. Les zadistes de Notre-Dame-des-Landes avaient un intérêt qui n'était pas monétaire contrairement à leurs adversaires. Leur comportement maximise une fonction que nous sommes bien incapables de décrire en termes numériques. Il y a aussi les gestes altruistes. Qu'est-ce que je maximise lorsque je décide de donner deux euros à une sdf, que je ne reverrai plus jamais, à la sortie du métro ?

Bouyssou Denis et Philippe Vincke, Relations binaires et modélisation des préférences. (Révisé janvier 2003). 2005.

Briggs, R. A., Normative Theories of Rational Choice: Expected Utility, *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/spr2017/entries/rationality-normative-utility/

Luce R. Duncan et Howard Raiffa, *Games and Decisions*, Dover, New York, 1985.

Kahneman Daniel et Amos Tversky, Prospect theory: An analysis of decision under risk », 47, dans *Decision, Probability, and Utility*, Peter Gärdenfors and Nils-Eric Sahlin (dir.),1988, 183-214.

Von Neumann John et Oskar Morgenstern, *Theory of Games and Economic Behavior,*

Princeton University Press, Princeton, 1944.

Savage J. Leonard, *The Foundations of Statistics*, Dover, New York, 1972.

Shih En Lu, www.sfu.ca/~shihenl/.../3p%20Expected%20Utility%20Theory.pdf

Skyrms Brian, *Choice & Chance*, Wadsworth Publishing Company, Belmont California, 1986.

Tversky Amos, Intransitivity of preferences. *Psychological Review*, 76, 1969, 31-48.