Plus de mathématiques pour gagner plus

More mathematics to get more money in the "weakest link" game

Walid Ben-ameur

Electronic version

URL: <http://journals.openedition.org/msh/2971>

DOI: 10.4000/msh.2971

ISSN: 1950-6821

Publisher

Centre d'analyse et de mathématique sociales de l'EHESS

Printed version

Date of publication: 1 March 2006

ISSN: 0987-6936

Electronic reference

Walid Ben-ameur, « Plus de mathématiques pour gagner plus », *Mathématiques et sciences humaines* [Online], 173 | Printemps 2006, Online since 22 May 2006, connection on 23 July 2020. URL : <http://journals.openedition.org/msh/2971> ; DOI : <https://doi.org/10.4000/msh.2971>

PLUS DE MATHEMATIQUES POUR GAGNER PLUS AU « MAILLON FAIBLE »

Walid BEN-AMEUR¹

RÉSUMÉ – *Nous essayons dans cet article de combiner des outils mathématiques simples provenant des probabilités et de l'optimisation pour traiter un problème issu du jeu télévisé « le maillon faible ». Nous mettons en évidence des stratégies de jeu optimales dépendant des niveaux des joueurs, de leurs objectifs et de leur niveau de coopération.*

MOTS-CLÉS – Jeux, Optimisation, Probabilités, Programmation dynamique

SUMMARY – *More mathematics to get more money in the « weakest link » game. Some simple mathematical tools issued from probability and optimization theories are combined in this paper to study several problems related to the « the weakest link » TV game. We highlight some optimal game strategies depending on the players performance levels, their goals and their cooperation level.*

KEYWORDS – Game, Optimization, Probabilities, Dynamic Programming

1. INTRODUCTION

Bien qu'elles soient rigoureuses et parfois complexes, les sciences mathématiques ont gardé à travers les siècles un aspect ludique indéniable. Certaines branches des mathématiques telles que les probabilités ou la théorie des graphes ont démarré sous forme de jeux ou de défis que se lançaient les mathématiciens entre eux. Plusieurs autres branches telles que l'arithmétique, la géométrie ou l'algèbre se sont enrichies grâce à l'effort de plusieurs mathématiciens qui cherchaient à modéliser certains jeux et à trouver les stratégies gagnantes [Cajori, 1999□Chabert *et al.*, 1994□Collette, 1973]. Nous pouvons citer à titre d'exemples les travaux à propos du jeu du solitaire, du Rubik's cube, du jeu de la vie, du jeu de dames, etc. [Avis, Deza, 2001□Criton, 1997□Delahaye, 1999□Joyner, 2002].

Nous nous intéressons dans cet article au jeu télévisé « le maillon faible qui est diffusé dans plusieurs pays. Nous tenterons de déterminer des stratégies optimales du jeu en fonction du niveau des joueurs, de leurs objectifs et de leur niveau de coopération. Nous ferons appel à un peu de probabilités et à la programmation dynamique qui est une technique classique d'optimisation [Bellman, 2003□Sniedovich, 1992].

¹ Institut national des télécommunications, 9, rue Charles Fourier, 91011 Evry, walid.benameur@int-evry.fr

2. PRINCIPES DU JEU

Un groupe de joueurs de niveaux différents répondent aux questions posées par l'animatrice en fonction de leurs connaissances et tentent d'accumuler des sommes d'argent qui sont remises à la fin au gagnant du jeu. Les questions sont posées dans un ordre circulaire.

Avant d'entendre la question qui lui sera posée, chaque joueur a la possibilité de transformer une chaîne de bonnes réponses de ses prédécesseurs, lorsqu'elle existe, en argent qui sera additionné aux gains du groupe. Plus la chaîne de bonnes réponses est longue, plus la somme créditée est importante. Cependant, si par excès de confiance un joueur décide d'essayer de prolonger la chaîne, une mauvaise réponse de sa part détruirait complètement la chaîne de bonnes réponses et priverait le groupe de tout l'argent qu'aurait pu rapporter cette chaîne de bonnes réponses. D'un autre côté, si les joueurs transforment trop vite les chaînes de bonnes réponses en argent, les chaînes de bonnes réponses deviennent courtes et rapportent peu d'argent. « Banquer » est le terme utilisé dans le jeu pour désigner le fait de transformer une chaîne de bonnes réponses en argent.

Pour résumer, le fait de banquer ou de mal répondre à une question remet à 0 le compteur des bonnes réponses consécutives. Le but du jeu est *a priori* de maximiser la somme des gains accumulés grâce aux chaînes de bonnes réponses.

Le jeu est en réalité composé de plusieurs manches. Un vote est organisé à la fin de chaque manche pour éliminer le joueur qui aurait été considéré par les autres comme étant « le maillon faible ». La toute dernière manche consiste en un affrontement entre les deux rescapés du jeu. Bien que les joueurs les moins efficaces soient généralement éliminés pendant les premières manches, la situation est parfois différente pendant les dernières manches. En effet, lorsqu'on se rapproche de la fin du jeu, certains joueurs craignant de perdre face à un joueur performant, votent contre les bons joueurs.

L'objectif de cette note est de proposer et étudier des stratégies de jeu permettant de maximiser les gains au cours d'une manche en banquant au « bon moment » grâce à de simples calculs probabilistes combinés à la programmation dynamique.

Nous adoptons les mêmes valeurs de gain en fonction de la longueur de la chaîne de bonnes réponses que celles qui sont réellement utilisées dans l'émission télévisée :

0	1	2	3	4	5	6	7	8	9
0	50	100	200	400	1000	2000	3000	4000	5000

Tableau 1. Gains en euros en fonction de la longueur de la chaîne de bonnes réponses

Ainsi, si les joueurs viennent de répondre correctement à 5 questions consécutives et si le joueur suivant décide de banquer avant de répondre, alors une somme de 1000 euros est automatiquement créditée au compte et on recommence à répondre aux questions et à essayer de former des chaînes de bonnes réponses.

L'élément-clé qui va conditionner notre stratégie de jeu est la probabilité de bonne réponse. Supposons par exemple que 30 questions sont posées au cours d'une partie de maillon faible. Alors, si tous les joueurs répondent correctement à toutes les questions, il sera possible de gagner 15200 euros en banquant au bon moment (après chaque

9 bonnes réponses pendant les 27 premières questions, et après les 3 dernières bonnes réponses). Cependant, la probabilité de bien répondre est rarement égale à 1. Que doit-on faire alors pour maximiser les gains même lorsque la probabilité de bonne réponse est moins élevée?

Notons que nous allons dans cet article nous intéresser exclusivement à la manière de banker en supposant connues les probabilités de bonne réponse de chaque joueur. En d'autres termes, nous ne cherchons pas à indiquer à un joueur donné s'il doit bien ou mal répondre aux questions.

3. UNE STRATEGIE OPTIMALE POUR MAXIMISER SES GAINS AU «MAILLON FAIBLE»

Plaçons-nous dans le cas où le nombre de questions posées est connu. On le désignera par N . Appelons $V[k]$ la somme d'argent gagnée (bankée) lorsqu'une chaîne de bonnes réponses de longueur k est transformée en argent. On supposera également que les joueurs répondent correctement aux questions avec une certaine probabilité p . On admettra dans un premier temps que la probabilité est la même pour tous les joueurs. Le cas plus général sera traité dans la partie 6.

Le nombre N , le vecteur des gains $V[k]$ (k variant de 0 à 9) et la probabilité p sont donc des données du problème. Les valeurs $V[k]$ seront celles du Tableau 1.

Pendant le déroulement du jeu, nous dirons que nous sommes à l'instant t (t variant de 0 à N) si exactement t questions ont déjà été posées. Désignons par $Gain[t, k]$ la moyenne des gains qu'on peut espérer avoir à partir de l'instant t sachant qu'il y a une chaîne de bonnes réponses de longueur k , étant entendu que nous appliquerons la meilleure stratégie possible. Les nombres $Gain[t, k]$ sont pour le moment inconnus et nous nous proposons de les calculer en déterminant par la même occasion la stratégie optimale de jeu qui permettrait de maximiser l'espérance des gains. Il est important de comprendre que $Gain[t, k]$ représente ce qu'on espère gagner à partir de l'instant t sachant qu'il y a déjà k bonnes réponses consécutives et non pas ce qui a déjà été gagné. Puisqu'en démarrant le jeu nous nous trouvons à l'instant $t = 0$ et que nous avons également $k = 0$, $Gain[0, 0]$ représente la somme qu'on peut espérer gagner en moyenne après la fin du jeu sachant que nous allons appliquer la meilleure stratégie possible.

Pour t inférieur ou égal à N , on établit facilement les relations suivantes:

$$\overline{Gain[t, k]} = \begin{cases} \text{Maximum} \begin{cases} V[k] + p \overline{Gain[t+1, 1]} + (1-p) \overline{Gain[t+1, 0]} \\ p \overline{Gain[t+1, k+1]} + (1-p) \overline{Gain[t+1, 0]} \end{cases} & \text{si } k < 9 \text{ et } t < N \\ V[k] + p \overline{Gain[t+1, 1]} + (1-p) \overline{Gain[t+1, 0]} & \text{si } k = 9 \text{ et } t < N \\ V[k] & \text{si } t = N \end{cases}$$

En effet, pour $k < 9$ et $t < N$, si le joueur banque, alors une somme $V[k]$ est créditée au compte du groupe. S'il répond ensuite correctement, le joueur suivant se retrouvera avec une chaîne contenant exactement 1 bonne réponse et on peut par définition espérer gagner $Gain[t+1, 1]$ à partir de ce moment. En revanche, si le joueur

courant se trompe, le suivant se retrouvera avec une chaîne vide qui ne permet d'espérer de gagner que $Gain[t+1,0]$ à partir de ce moment et jusqu'à la fin de la manche. En d'autres termes, si le joueur actuel banque, l'espérance du gain à partir de ce moment est donnée par $V[k] + p \square Gain[t+1,1] + (1 \square p) \square Gain[t+1,0]$. La situation est différente si le joueur courant décide de ne pas banker $\square$ l'espérance du gain est donnée par $p \square Gain[t+1,k+1] + (1 \square p) \square Gain[t+1,0]$. Comme nous cherchons à maximiser l'espérance du gain, la stratégie choisie sera celle qui correspond au maximum de $V[k] + p \square Gain[t+1,1] + (1 \square p) \square Gain[t+1,0]$ et

$$p \square \overline{Gain[t+1,k+1]} + (1 \square p) \square \overline{Gain[t+1,0]}.$$

Ainsi pour calculer les nombres $\overline{Gain[t,k]}$, il suffit de procéder par programmation dynamique en commençant avec $t = N$ et en diminuant t progressivement jusqu'à arriver à 0. La formule de récurrence établie nous permet de procéder ainsi.

La stratégie optimale est déduite directement du calcul précédent. En effet, si dans le calcul de $Gain[t,k]$ le maximum est atteint avec $V[k] + p \square Gain[t+1,1] + (1 \square p) \square Gain[t+1,0]$, alors il faut banker si on se trouve à l'instant t avec une chaîne de bonnes réponses de longueur k . Dans le cas contraire, on ne doit pas banker.

Pour illustrer nos propos, nous indiquons dans le Tableau 2 la stratégie à suivre (1 pour banker, 0 pour ne pas banker et $-$ lorsque la chaîne de bonnes réponses est de longueur 0) ainsi que le gain moyen qu'on pourrait espérer en suivant cette stratégie optimale. Pour cet exemple, nous avons pris $N = 4$ et nous avons considéré deux valeurs différentes des probabilités de bonne réponse $p = 0,8$ et $p = 0,7$. Nous avons adopté les gains définis dans le Tableau 1. Ainsi si les joueurs estiment qu'ils ont une probabilité $p = 0,8$ de bien répondre, alors ils doivent regarder dans le Tableau 2 la stratégie correspondant à la situation dans la quelle ils se trouvent $\square$ à chaque valeur de t et de k nous indiquons ce qu'il faut faire. On peut observer que les stratégies sont les mêmes pour les deux valeurs de p sauf lorsque $t = 2$ et $k = 1$. On constate également une diminution importante du gain moyen qui passe de 206,72 à 147,385 lorsque la probabilité de bonne réponse passe de 0,8 à 0,7.

	$p = 0,8$		$p = 0,7$	
$t = 0 \ k = 0$	Strategie = -	Gain = 206,72	Strategie = -	Gain = 147,385
$t = 1 \ k = 0$	Strategie = -	Gain = 124,8	Strategie = -	Gain = 105
$t = 1 \ k = 1$	Strategie = 0	Gain = 227,2	Strategie = 0	Gain = 165,55
$t = 2 \ k = 0$	Strategie = -	Gain = 80	Strategie = -	Gain = 70
$t = 2 \ k = 1$	Strategie = 0	Gain = 136	Strategie = 1	Gain = 120
$t = 2 \ k = 2$	Strategie = 0	Gain = 264	Strategie = 0	Gain = 206,5
$t = 3 \ k = 0$	Strategie = -	Gain = 40	Strategie = -	Gain = 35
$t = 3 \ k = 1$	Strategie = 1	Gain = 90	Strategie = 1	Gain = 85
$t = 3 \ k = 2$	Strategie = 0	Gain = 160	Strategie = 0	Gain = 140
$t = 3 \ k = 3$	Strategie = 0	Gain = 320	Strategie = 0	Gain = 280
$t = 4 \ k = 0$	Strategie = -	Gain = 0	Strategie = -	Gain = 0
$t = 4 \ k = 1$	Strategie = 1	Gain = 50	Strategie = 1	Gain = 50
$t = 4 \ k = 2$	Strategie = 1	Gain = 100	Strategie = 1	Gain = 100
$t = 4 \ k = 3$	Strategie = 1	Gain = 200	Strategie = 1	Gain = 200
$t = 4 \ k = 4$	Strategie = 1	Gain = 400	Strategie = 1	Gain = 400

Tableau 2. La stratégie optimale pour un nombre de questions posées $N = 4$

Notons que les valeurs $\overline{Gain[0,0]}$ du Tableau 2 ne sont pas des sommes garanties : il s'agit seulement de la moyenne des gains qu'on peut avoir si on applique rigoureusement la stratégie optimale qu'on vient de définir pour une probabilité donnée de bonne réponse.

Une illustration plus complète est proposée dans la Figure 1 où on a considéré le cas $N = 30$ avec des probabilités de bonnes réponses allant de 0 à 1. Une des courbes de la Figure 1, intitulée optimum, représente le gain moyen en fonction de la probabilité de bonne réponse. Nous observons que le gain qu'on peut espérer avoir en appliquant une stratégie optimale reste relativement faible pour des probabilités de bonnes réponses inférieures à approximativement 0,57. L'augmentation du gain est linéaire en fonction de p lorsque p est inférieur à cette valeur. Les gains augmentent ensuite de plus en plus vite avec la probabilité de bonne réponse. Ceci revient aussi à dire que pour des probabilités élevées de bonnes réponses, une légère baisse de cette probabilité a des répercussions importantes sur le montant des gains. Ainsi le gain moyen est de l'ordre de 8867,54 euros pour $p = 0,9$ alors qu'il descend à 4544,80 euros pour $p = 0,8$ et il ne dépasse pas 1143,43 euros lorsque $p = 0,6$.

Notons que la valeur charnière $p \approx 0,57$ dépend de la valeur de N . Elle correspond à la plus petite valeur de p à partir de laquelle on n'a plus nécessairement intérêt à banker dès qu'il y a une bonne réponse. En d'autres termes, à partir de $p \approx 0,57$ le nombre $p \square Gain[t+1, k+1] + (1 \square p) \square Gain[t+1, 0]$ commence à dépasser $V[k] + p \square Gain[t+1, 1] + (1 \square p) \square Gain[t+1, 0]$ pour certaines valeurs de t et de k .

Figure 1

Nous observons également une discontinuité de la dérivée aux alentours de la même valeur $p \approx 0,57$. Le même phénomène est constaté pour les autres valeurs de N (cf. par exemple la Figure 3 où $N = 4$). Ces discontinuités proviennent de la nécessité de changer radicalement de stratégie de jeu lorsque la probabilité dépasse un certain seuil. Le paragraphe suivant propose un autre éclairage de la situation. Notons également qu'il est facile de voir grâce aux formules de récurrence déjà exposées que $\text{Gain}[0,0]$ est une fonction en p qui est continue et polynomiale par morceaux.

Nous représentons également sur la Figure 2 les stratégies optimales à adopter pour un nombre de questions $N = 30$ et pour quatre valeurs différentes de p . L'axe des abscisses indique le numéro de la question t et l'axe des ordonnées représente la longueur de la chaîne de bonnes réponses k . Nous indiquons sur les quatre graphes les cas où il faut banker pour maximiser le gain moyen. Ces stratégies ont évidemment été obtenues en utilisant les relations de récurrence précédemment établies.

On peut par exemple observer qu'à l'instant $t = 20$, on a intérêt à banker quelque soit la longueur de la chaîne de bonnes réponses lorsque $p = 0,1$. En revanche, on voit que dans le cas où $p = 0,5$, on ne doit pas banker si la chaîne de bonnes réponses est de longueur 3 ou 4. Lorsque p passe à 0,6, on doit s'abstenir de banker tant que k est strictement inférieur à 6. Enfin pour $p = 0,8$, on ne banque que si $k = 8$ ou 9. Remarquons aussi que la valeur approximative $p \approx 0,57$ évoquée précédemment se trouve entre 0,5 et 0,6 et on voit bien que la stratégie de jeu optimale correspondant à $p = 0,6$ est bien différente de celle qui consiste à banker tout de suite.

Figure 2

4. STRATÉGIE ÉLÉMENTAIRE BANQUER APRES x BONNES RÉPONSES

Pour mieux comprendre la nature de la stratégie optimale (décrite précédemment et calculée par programmation dynamique), nous avons cherché à la comparer à d'autres stratégies intuitives et plus élémentaires. Pour un nombre x donné entre 1 et 9, nous considérons la stratégie de jeu qui consiste à ne banker qu'après l'obtention de x bonnes réponses successives. Une exception triviale est faite après la dernière question du jeu où on banque même si on n'a pas atteint x bonnes réponses pour éviter de perdre bêtement de l'argent. Le même genre de formules de récurrence permet de calculer le gain obtenu par une telle stratégie. Nous notons $\overline{Gain}_x[t, k]$ le gain qu'on peut espérer avoir si on se trouve à l'instant t avec k bonnes réponses successives ($k \leq x, k \leq 9$) en appliquant la stratégie qu'on vient de décrire.

Pour t inférieur ou égal à N , on établit facilement la relation

$$\overline{Gain}_x[t, k] = \begin{cases} p \overline{Gain}_x[t+1, k+1] + (1-p) \overline{Gain}_x[t+1, 0] & \text{si } k < x \text{ et } t < N \\ V[x] + p \overline{Gain}_x[t+1, 1] + (1-p) \overline{Gain}_x[t+1, 0] & \text{si } k = x \text{ et } t < N \\ V[k] & \text{si } t = N \end{cases}$$

La Figure 1 contient trois courbes supplémentaires correspondant à $x = 1$, $x = 5$ et $x = 9$, qui représentent les gains moyens qu'on peut espérer obtenir en appliquant les stratégies qu'on vient de décrire. En comparant ces trois courbes avec la courbe correspondant à la stratégie optimale, nous observons évidemment que le gain obtenu est inférieur ou égal à ce que donnerait la stratégie optimale. On voit également qu'il n'y a aucune stratégie du type banker après x bonnes réponses qui soit toujours optimale plus la stratégie est bonne pour les faibles valeurs de p , plus elle sera mauvaise pour les fortes valeurs de p et *vice versa*. On voit ainsi que la stratégie correspondant à $x = 1$ est optimale pour des valeurs de p inférieures approximativement à 0,57 mais elle devient rapidement très mauvaise pour des valeurs plus élevées de p . Nous constatons exactement l'inverse pour $x = 9$.

5. L'INTÉRÊT DE BIEN CONNAÎTRE SES MOYENS

La stratégie optimale qui a été présentée dépend directement de la probabilité de bonne réponse p . Surestimer ou sous-estimer cette probabilité peut faire perdre de l'argent. Supposons par exemple que la probabilité réelle de bonne réponse est $p = 0,7$ alors que les joueurs pensaient qu'elle serait de l'ordre de 0,9. Les joueurs appliqueront donc la stratégie optimale correspondant à $p = 0,9$. Cette stratégie est pratiquement confondue avec celle qui consiste à banker après $x = 9$ bonnes réponses (cf. Figure 1). Cependant, on voit que pour $p = 0,7$ le gain moyen de cette stratégie est de l'ordre de 1864,12 alors que la stratégie optimale permet d'avoir un gain moyen de l'ordre de 2296,15. Il y a ainsi une perte moyenne de l'ordre de 432 à la suite de la surestimation de la probabilité de bonne réponse. De même si les joueurs sous-estiment la probabilité de bonne réponse en pensant qu'elle ne vaut que 0,5 alors ils devront appliquer la stratégie optimale correspondant à $p = 0,5$ qui consiste à banker après $x = 1$ bonne réponse (cf. Figure 1). Le gain moyen qu'ils pourront obtenir avec cette stratégie est de

l'ordre de 1050, ce qui est largement en dessous de ce qu'ils auraient gagné avec la stratégie optimale correspondant à $p = 0,7$ (2296,15).

6. PROBABILITÉS VARIABLES

Nous avons supposé dans tout ce qui précède que les joueurs avaient tous le même niveau de performance. En réalité, ils n'ont pas nécessairement la même probabilité de bonne réponse. Dans ce cas, on a tout intérêt à donner pour chaque joueur une bonne estimation de sa propre probabilité de bonne réponse et d'établir une stratégie optimale en tenant compte de ces paramètres. Le fait que les joueurs soient hétérogènes se traduit par différentes probabilités $p[t]$ pour répondre à la question t ($0 \leq t \leq N$). Les formules de récurrence permettant d'établir la stratégie optimale deviennent

$$\overline{Gain}[t, k] = \begin{cases} \text{Maximum} \begin{cases} V[k] + p[t+1] \overline{Gain}[t+1, 1] + (1 - p[t+1]) \overline{Gain}[t+1, 0] \\ p[t+1] \overline{Gain}[t+1, k+1] + (1 - p[t+1]) \overline{Gain}[t+1, 0] \end{cases} & \text{si } k < 9 \text{ et } t < N \\ V[k] + p[t+1] \overline{Gain}[t+1, 1] + (1 - p[t+1]) \overline{Gain}[t+1, 0] & \text{si } k = 9 \text{ et } t < N \\ V[k] & \text{si } t = N \end{cases}$$

Remarquons qu'il est également possible de tenir compte de probabilités dynamiques de bonnes réponses. Imaginons par exemple que plus la longueur de la chaîne de bonnes réponses est longue plus les joueurs ont tendance à commettre des erreurs («facteur de stress»). La probabilité de bonne réponse serait donc une fonction de k et de t ($p[t, k]$). Le même type de formules de récurrence nous permet d'établir encore une fois la stratégie optimale.

Nous avons mentionné dans la partie 2 que le jeu comprenait plusieurs manches qui pouvaient avoir des difficultés différentes. On peut donc s'attendre à ce que la performance des joueurs dépende de la manche. Cette variabilité peut être intégrée en considérant des probabilités dépendant, entre autres, de la manche.

Un autre aspect important du jeu est l'élimination d'un joueur à la fin de chaque manche. Nous avons déjà observé que le niveau d'un joueur pouvait avoir une incidence sur le vote des autres joueurs qui peuvent parfois l'éliminer par crainte de l'affronter à la dernière manche. Pour tenir compte de cette particularité, on peut admettre qu'un joueur puisse donner de temps en temps une mauvaise réponse à une question alors qu'il connaissait la bonne réponse. Ceci se modélise simplement par une probabilité de bonne réponse plus faible pour ce joueur. On peut aussi supposer qu'un joueur qui décide à l'avance de mal répondre, peut malgré cela banker en utilisant une stratégie optimale. En effet, la dernière manche du jeu se base seulement sur le niveau des deux joueurs restant et non pas sur leur façon de banker. Il n'y a donc pas de raison apparente pour qu'un joueur décide de mal banker puisque c'est seulement son niveau qui peut faire peur aux autres joueurs. Nous verrons cependant dans la partie 8 comment tenir compte des joueurs qui ont une autre stratégie de jeu.

Sous les conditions qu'on vient de décrire, les probabilités $p[t, k]$ qui doivent être utilisées pour déterminer la stratégie optimale par programmation dynamique, sont des estimations basées non seulement sur le niveau des joueurs mais intègrent également

leur volonté de ne pas se faire éliminer par les autres. L'estimation de ces probabilités est un problème intéressant en soi qui pourrait faire intervenir des modèles statistiques élaborés. L'objectif de cet article étant d'étudier les stratégies de banker et non pas les stratégies de réponses aux questions, nous n'allons pas nous attarder sur ce problème d'estimation des probabilités de bonnes réponses.

7. ET SI L'ON CONNAISSAIT L'AVENIR ?

Pour mesurer encore plus l'intérêt d'adopter une stratégie maximisant le gain moyen, nous allons la comparer avec la stratégie «*inictive*» qui suppose que l'on connaît à l'avance les questions qui recevront une bonne réponse et celles qui recevront une mauvaise réponse.

En effet, on pourrait dans ce cas déterminer sans ambiguïté à quel moment doivent banker les joueurs pour maximiser les gains. Cette stratégie est évidemment purement imaginaire mais elle nous donne une borne supérieure qu'il est impossible de dépasser quelle que soit la stratégie de jeu utilisée. Nous supposons ici pour simplifier que les joueurs sont homogènes et nous désignons toujours par p la probabilité de bonne réponse. Etant donné p , nous considérons toutes les réalisations possibles du jeu. Si k questions parmi N reçoivent une bonne réponse, cette réalisation se produit avec une probabilité de $p^k (1-p)^{N-k}$ si on suppose que les réponses aux questions sont indépendantes entre elles. Nous calculons assez facilement pour chaque réalisation le gain maximal qu'on peut avoir en bankant au bon moment. Ce gain est obtenu en rallongeant au maximum (lorsque $k < 9$) les chaînes de bonnes réponses et en bankant juste avant de commettre une erreur.

Ensuite nous calculons la moyenne de ces gains en tenant compte de la probabilité de chaque réalisation (i.e., $p^k (1-p)^{N-k}$ où k est le nombre de bonnes réponses). La moyenne obtenue représente donc le gain moyen maximal qu'on peut avoir *a posteriori* pour une probabilité de bonne réponse p . Le terme «*a posteriori*» provient du fait que nous supposons ici connaître la réalisation du jeu. Notons que pour calculer cette moyenne nous avons besoin de considérer de l'ordre de 2^N événements probabilistes possibles ce qui peut être prohibitif en terme de calcul. Nous nous contenterons donc du cas $N = 4$ pour illustrer notre discours.

Nous observons sur la Figure 3 une première courbe représentant le gain moyen qu'on peut obtenir en appliquant la stratégie optimale. Une seconde courbe représente la moyenne du gain maximal qu'on peut avoir *a posteriori*. La seconde courbe est évidemment au-dessus de la première mais la différence entre les deux courbes reste limitée. Notons cependant que la différence entre les deux courbes devrait dépendre du nombre de questions posées N . Plus d'expérimentations numériques sont nécessaires pour tirer des résultats généraux sur l'écart entre le gain obtenu par une stratégie optimale et le gain moyen *a posteriori*. Signalons enfin que cet écart entre les deux courbes constitue ce qu'on peut appeler le prix de l'incertitude ce qu'on perd en moyenne lorsqu'on connaît le niveau des joueurs (p) sans savoir quelles sont les questions qui recevront une bonne réponse.

Figure 3

8. ET SI CERTAINS JOUEURS NE COOPÉRAIENT PAS

Connaissant la probabilité de bonne réponse p (si les joueurs sont homogènes) ou les probabilités $p[t]$ (si les joueurs sont hétérogènes), le gain moyen qu'on peut obtenir en appliquant la stratégie optimale est fondé sur la coopération de tous les joueurs qui doivent appliquer la même stratégie quand ça sera à eux de jouer. Cependant, il se peut que certains joueurs ne soient pas coopératifs par exemple parce qu'il ne font pas confiance aux mathématiques. Que doivent faire alors les autres joueurs pour limiter les dégâts et maximiser le gain moyen en tenant compte des paramètres du jeu. La réponse semble difficile si on ne connaît rien sur la stratégie qui sera adoptée par les joueurs non coopératifs. Il est en revanche facile d'établir une stratégie optimale si on sait ce que vont faire les joueurs non coopératifs. Supposons par exemple que le joueur qui répond à la question numéro $t_0 + 1$ et qui va donc décider à l'instant t_0 s'il faut banker ou pas, est un joueur non coopératif qui banque d'une façon automatique dès qu'il y a une bonne réponse. Dans ce cas les formules de récurrence précédentes restent valables pour tout $t \neq t_0$ et on a pour $t = t_0$

$$\overline{Gain[t_0, k]} = V[k] + p[t_0 + 1] \overline{Gain[t_0 + 1, 1]} + (1 - p[t_0 + 1]) \overline{Gain[t_0 + 1, 0]}$$

Ainsi la stratégie optimale s'établit en tenant compte des décisions des joueurs non coopératifs. Un exemple est montré au Tableau 3 où on a considéré les mêmes paramètres qu'au Tableau 2 tout en supposant qu'un joueur non coopératif banque automatiquement entre la troisième et la quatrième question. En comparant les deux tableaux, nous constatons évidemment une diminution du gain moyen lorsqu'on a des

joueurs non coopératifs qui passe de 206,72 à 164,8 pour $p = 0,8$ et de 147,385 à 140 pour $p = 0,7$. Nous constatons également que le fait qu'un joueur non coopératif banque après la troisième question modifie la stratégie optimale ($p = 0,8$) puisqu'il faut maintenant banker même si on se trouve à l'état $t = 2, k = 1$ alors qu'on ne le faisait pas dans le premier cas (sans joueurs non coopératifs). De même pour $p = 0,7$ on voit qu'il faut maintenant banker si $t = 1, k = 1$ alors que ceci n'était pas recommandé auparavant.

	$p = 0,8$		$p = 0,7$	
$t = 0, k = 0$	Strategie = -	Gain = 164,8	Strategie = -	Gain = 140
$t = 1, k = 0$	Strategie = -	Gain = 120	Strategie = -	Gain = 105
$t = 1, k = 1$	Strategie = 0	Gain = 176	Strategie = 1	Gain = 155
$t = 2, k = 0$	Strategie = -	Gain = 80	Strategie = -	Gain = 70
$t = 2, k = 1$	Strategie = 1	Gain = 130	Strategie = 1	Gain = 120
$t = 2, k = 2$	Strategie = 0	Gain = 200	Strategie = 0	Gain = 175
$t = 3, k = 0$	Strategie = -	Gain = 40	Strategie = -	Gain = 35
$t = 3, k = 1$	Strategie = 1	Gain = 90	Strategie = 1	Gain = 85
$t = 3, k = 2$	Strategie = 1	Gain = 140	Strategie = 1	Gain = 135
$t = 3, k = 3$	Strategie = 1	Gain = 240	Strategie = 1	Gain = 235
$t = 4, k = 0$	Strategie = -	Gain = 0	Strategie = -	Gain = 0
$t = 4, k = 1$	Strategie = 1	Gain = 50	Strategie = 1	Gain = 50
$t = 4, k = 2$	Strategie = 1	Gain = 100	Strategie = 1	Gain = 100
$t = 4, k = 3$	Strategie = 1	Gain = 200	Strategie = 1	Gain = 200
$t = 4, k = 4$	Strategie = 1	Gain = 400	Strategie = 1	Gain = 400

Tableau 3. La stratégie optimale pour un nombre de questions posées $N = 4$ sachant qu'un joueur non coopératif banque automatiquement après la troisième question

9. UN AUTRE OBJECTIF IMPLIQUE UNE AUTRE STRATÉGIE DÉPASSER UN CERTAIN SEUIL SUR L'ENSEMBLE DE LA MANCHE

La stratégie optimale développée précédemment est basée sur l'espérance du gain qu'on cherche à maximiser. Bien que très intuitif, ce critère n'est pas unique. Les joueurs peuvent par exemple s'estimer satisfaits à partir du moment où ils gagnent plus qu'une certaine somme. Étant donné un seuil θ , l'objectif des joueurs consisterait à maximiser la probabilité de gagner plus que θ . Pour établir la stratégie optimale qui permettrait de maximiser cette probabilité, nous introduisons les variables $q_\theta[t, k, \theta]$ représentant la probabilité de dépasser le seuil θ sachant qu'à l'instant t nous disposons d'une chaîne de bonnes réponses de longueur k et d'une somme gagnée dans le passé égale à un nombre entier θ . Si θ est plus grand que θ , on aura évidemment $q_\theta[t, k, \theta] = 1$. La stratégie optimale s'établit grâce aux formules de récurrence suivantes valables pour $\theta < \theta$

Figure 4

$$\begin{aligned}
 & \begin{cases} 1 & \text{si } t + V[k] \geq \square \\ 0 & \text{si } t = N \text{ et } t + V[k] < \square \end{cases} \\
 q_{\square}[t, k, \square] = & \begin{cases} \text{Max} \begin{cases} p[t+1]q_{\square}[t+1, 1, \square + V[k]] + (1-p[t+1])q_{\square}[t+1, 0, \square + V[k]] \\ p[t+1]q_{\square}[t+1, k+1, \square] + (1-p[t+1])q_{\square}[t+1, 0, \square] \end{cases} & \text{si } k < 9, t < N \text{ et } t + V[k] < \square \\ p[t+1]q_{\square}[t+1, 1, \square + V[k]] + (1-p[t+1])q_{\square}[t+1, 0, \square + V[k]] & \text{si } k = 9, t < N \text{ et } t + V[k] < \square \end{cases}
 \end{aligned}$$

Le calcul des probabilités se fait comme précédemment par programmation dynamique en exploitant les relations de récurrence qu'on vient de présenter.

Pour illustrer nos propos, nous reprenons l'exemple de la Figure 1 où le nombre de questions posées N est égal à 30 et nous représentons sur la Figure 4 la probabilité de dépasser un seuil valant 1000 ou 10000 euros en fonction de la probabilité de bonne réponse (on suppose ici que les joueurs ont des niveaux homogènes). On observe évidemment que la probabilité de dépasser le seuil diminue lorsque le seuil augmente ou lorsque la probabilité de bonne réponse diminue. Notons que les courbes de la Figure 4 ont été obtenues en appliquant pour chaque seuil et chaque probabilité de bonne réponse la stratégie optimale leur correspondant. On ne joue pas de la même façon lorsqu'on cherche à gagner 1000 ou 10000 euros.

Nous représentons également sur la Figure 4 la probabilité de dépasser chacun des seuils définis précédemment en appliquant la stratégie maximisant l'espérance des gains. Pour calculer ces probabilités, nous procédons comme suit pour chaque

probabilité de bonne réponse, nous déterminons la stratégie qui maximise l'espérance des gains en utilisant l'algorithme de la première section. Ensuite, nous calculons, pour chaque probabilité de bonne réponse, la probabilité de dépasser un certain seuil en procédant toujours par programmation dynamique et en considérant la stratégie maximisant l'espérance qu'on vient de déterminer.

Les résultats reportés sur la Figure 4 montrent un résultat surprenant : la probabilité de dépasser un certain seuil peut diminuer avec la probabilité de bonne réponse. Ce résultat est dû au fait que nous appliquons ici la stratégie qui maximise l'espérance et non pas celle qui maximise la probabilité de dépasser un seuil.

10. QUE FAIRE SI LE GAIN DANS UNE MANCHE EST LIMITE D'UNE MANIÈRE AUTORITAIRE

Pour finir notre étude, nous allons prendre en compte un autre paramètre qui est une limite supérieure sur le gain que peuvent avoir les joueurs pendant une manche du jeu. En effet, les concepteurs du jeu peuvent introduire une telle limite pour éviter de déboursier des sommes « trop » importantes. La limite fixée dans la version télévisée du jeu semble être 5000 euros par manche. Appelons cette limite supérieure L et considérons les nombres $Gain_L[t, k, \square]$ représentant l'espérance du gain futur si on se trouve à l'instant t avec k bonnes réponses successives et un gain déjà acquis de valeur $\square$ sachant d'une part que le gain total ne peut dépasser L et d'autre part que l'on va appliquer la meilleure stratégie possible. Il est aisé de voir que les nombres vérifient les relations de récurrence suivantes ($t \leq N, k \leq 9, \square \leq L$) :

$$Gain_L[t, k, \square] = \begin{cases} L & \text{si } \square + V[k] \geq L \\ \square + V[k] & \text{si } t = N \text{ et } \square + V[k] < L \\ \max \begin{cases} p[t+1] Gain_L[t+1, 1, \square + V[k]] + (1-p[t+1]) Gain_L[t+1, 0, \square + V[k]] \\ p[t+1] Gain_L[t+1, k+1, \square] + (1-p[t+1]) Gain_L[t+1, 0, \square] \\ p[t+1] Gain_L[t+1, 1, \square + V[k]] + (1-p[t+1]) Gain_L[t+1, 0, \square + V[k]] \end{cases} & \text{si } k < 9, t < N \text{ et } \square + V[k] < L \end{cases}$$

Pour comprendre l'effet de cette limitation des gains sur la stratégie adoptée et sur le montant moyen des gains, nous avons tracé sur la Figure 5 une première courbe représentant le gain moyen lorsqu'il n'y a aucune limitation (la même que la courbe « Optimum » de la Figure 1). Nous avons également représenté le gain moyen avec une limitation à 3000 et une limitation à 5000 euros. Nous observons que plus la limitation est forte plus le gain moyen diminue. Cependant, pour des faibles valeurs de p , on voit que les trois courbes sont presque confondues. On observe un certain écart entre les courbes lorsque la probabilité de bonne réponse augmente. Il est aussi important de voir que même lorsque le gain moyen de la stratégie optimale (sans limitation) est inférieur à 5000 euros (par exemple pour $p = 0,8$, le gain moyen s'élève à 4544,8), le fait d'introduire une limitation du gain à 5000 euros fait baisser le gain moyen qui devient 3629,56 euros. Remarquons également que dans le cas « Sans limitation » on observe une croissance importante du gain lorsque p se rapproche de 1 alors que la situation est tout à fait différente dans le cas « Avec limitation » où on observe une stagnation du gain qui se rapproche lentement de la limite L .

11. CONCLUSION

Nous avons essayé dans cet article de présenter une première étude du jeu télévisé «*Le maillon faible*». Nous avons proposé une association d'outils mathématiques pour déterminer des stratégies de jeu optimales. Bien que plusieurs comportements possibles des joueurs et différents objectifs aient été pris en compte dans cette étude, nous pensons qu'elle est loin d'être complète parce que l'humain reste difficile à modéliser par des équations.

Figure 5

BIBLIOGRAPHIE

- AVIS D., DEZA A., "On the solitaire cone and its relationship to multicommodity flows", *Mathematical Programming*, Vol. 90, 2001, p. 27-57.
- BELLMAN R. E., *Dynamic programming*, New-York, Dover Publications, 2003.
- CAJORI F., *A history of mathematics*, AMS Chelsea, 1999.
- CHABERT J.-L. & al., *Histoires d'algorithmes. Du caillou à la puce*, Paris, Belin, 1994.
- COLLETTE J.-P., *Histoire des mathématiques*, Montréal, Éditions du renouveau pédagogique, 1973.

CRITON M., *Les jeux mathématiques*, Paris, Presses universitaires de France, 1997.

DELAHAYE J.-P., *Jeux mathématiques et mathématique des jeux*, Paris, Belin, 1999.

JOYNER D., *Adventures in group theory: Rubik's cube, Merlin's machine, and other mathematical toys*, Baltimore (Maryland), Johns Hopkins University Press, April 2002.

SNIEDOVICH M., *Dynamic programming*, New-York, Marcel Dekker, 1992.