

Stratégies optimales de recherche

Claude Loverdo

Résumé

Il existe beaucoup de systèmes qui présentent des problèmes de minimisation du temps de recherche. Le modèle des marches de Lévy a beaucoup été mis en avant ces dernières années. L'équipe qui est mon premier choix pour cette demande de bourse a développé un modèle alternatif de recherche intermittente, auquel j'ai contribué pendant mon stage de Master 2. Ma thèse présente trois axes : développements théoriques de de modèle de recherche intermittente ; application à la recherche par des enzymes de séquences-cibles sur l'ADN ; application aux trajectoires d'animaux.

1 Introduction : les problèmes de recherche

Pour introduire le sujet, on peut imaginer une situation type : la perte de clés sur une plage. Comment fait-on pour les retrouver au plus vite ? Quand on les cherche dans le sable, l'exploration de la surface est très lente. De plus, si on n'a pas de points de repère, si l'on n'adopte pas de stratégie systématique, on peut repasser sur des parcelles déjà fouillées. On peut de temps en temps se lever, faire quelques pas, et recommencer à chercher dans le sable (cf. Fig 1). En marchant, on ne peut pas trouver les clés, mais comme le déplacement est plus rapide, on peut mieux explorer l'espace. Cette phase de déplacement rapide mais sans possibilité de détecter la cible diminue-t-elle ou augmente-t-elle le temps mis à retrouver les clés ? Si ce type de stratégie est favorable, combien de temps faut-il passer dans chacune des deux phases pour minimiser le temps de recherche ?

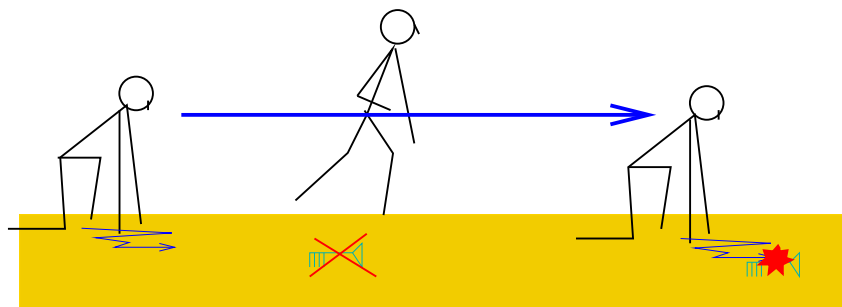


FIG. 1 – Chercheur de clé dans le sable. Soit il fouille le sable et avance peu, soit il marche rapidement mais ne peut pas détecter la clé.

Ce type de stratégie de recherche intermittente est rencontré dans de nombreuses situations. À l'échelle microscopique, des enzymes qui agissent sur des séquences spécifiques de l'ADN les trouvent très rapidement, et des expériences semblent indiquer un mélange de deux types de déplacement : de la diffusion lente le long de l'ADN, en une dimension, et de la diffusion en trois dimensions plus rapide [1]. À l'échelle macroscopique, des animaux à la recherche de différentes cibles (nourriture, partenaires, abris, etc. [2]) présentent des trajectoires avec une succession de déplacements rectilignes et d'arrêts suivis de réorientations. La vitesse dégrade la perception, donc la cible ne peut pas être détectée lors des déplacements. Inversement, pendant les arrêts, l'animal peut repérer sa cible dans une certaine aire autour de lui [3].

On s'attend à ce que le temps de recherche soit minimal. En effet dans les problèmes biologiques ci-dessus, bien que ni les animaux ni l'enzyme n'aient réfléchi consciemment à optimiser leur temps de recherche, si ce temps de recherche est un facteur limitant qui peut fortement influencer sur le succès reproductif de l'organisme, l'évolution a pu sélectionner des stratégies de recherche optimales.

Le modèle des marches de Lévy, largement mis en avant ces dernières années, notamment pour l'explication de trajectoires d'animaux ne prend pas du tout en compte cette intermittence, et la dégradation de la perception avec la vitesse. Dans ce modèle la trajectoire consiste en une série de réorientations instantanées et de segments droits, dont les longueurs suivent une loi de Lévy [4]. La probabilité d'avoir une nouvelle réorientation à l'instant t est proportionnelle à $t^{-\mu}$, avec $1 < \mu < 3$. La cible peut être détectée tout le long de la trajectoire. Quand le temps moyen de recherche est optimisé par rapport à μ [5], le μ optimum est 2 dans le cas de cibles revisitables, c'est-à-dire dans le cas où elles ne sont pas définitivement détruites après une première détection. Mais, dans le cas de cibles qui sont détruites, ce qui est aussi plus proche de problèmes de première détection d'une cible, $\mu_{optimum} \rightarrow 1$, ce qui donne une trajectoire tendant vers une ligne droite, ce qui est assez étonnant en regard de beaucoup de trajectoires réelles observées.

2 Modèle de recherche intermittente

Un modèle alternatif a été proposé par l'équipe constituée d'Olivier Bénichou, de Raphaël Voituriez, de Michel Moreau. J'y ai effectué mon stage de Master 2, et je souhaite y faire ma thèse (premier choix). L'idée principale de ce modèle est de tenir compte de cette intermittence observée, de cette alternance entre des phases lentes mais réactives (c'est-à-dire qui permettent de détecter la cible) et des phases rapides mais non-réactives (c'est-à-dire pendant lesquels la cible ne peut pas être détectée). Ce modèle a été originellement développé dans le cas de cibles régulièrement espacées à une dimension [6, 7, 8, 9].

J'ai contribué pendant mon stage de Master 2 à l'extension du modèle dans deux directions, qui ont donné lieu à trois articles [10, 11, 12].

D'une part, les cibles peuvent être placées de manière aléatoire, selon une distribution de Poisson. Cette distribution est intéressante car elle peut rendre compte du cas de cibles dont les positions sont indépendantes les unes des autres. Les calculs analytiques sont beaucoup plus difficiles que dans le cas de cibles régulières, mais des approximations, validées par des simulations numériques, montrent que l'intermittence est favorable. De plus, nous obtenons une relation entre le temps à passer dans la phase lente et réactive (décrite par un mouvement diffusif de coefficient de diffusion D), et le temps à passer dans la phase rapide et non-réactive (décrite par un mouvement balistique de vitesse v) [10].

D'autre part, un espace à deux dimensions semble plus réaliste pour certains des problèmes de recherche évoqués dans l'introduction. Là encore les calculs analytiques sont plus difficiles, même dans le cas de cibles régulièrement espacées. Cependant des approximations validées par des simulations numériques sont possibles. La phase rapide est décrite par un mouvement balistique de vitesse v avec une orientation aléatoire. La phase lente est décrite par une diffusion de coefficient D avec détection immédiate si la cible est à une distance inférieure à a , ou par un arrêt avec un taux de détection de la cible par unité de temps k si la cible est à une distance inférieure à a . Dans les deux cas, on obtient un optimum pour un couple unique de durées moyennes des deux phases. De plus, nous obtenons que la durée optimale de la phase rapide non-réactive est la même dans les deux cas, indépendamment de la description de la phase lente réactive. Cela montre la robustesse du modèle [11, 12].

Dans ce domaine de recherche qui connaît une grande activité ces dernières années [4, 5, 13, 14], ce modèle qui est une alternative convaincante au modèle des marches de Lévy a particulièrement éveillé l'attention, puisqu'il est même largement commenté dans un très récent *News and views of Nature* [15].

Ma thèse s’articule autour de trois axes principaux : continuer à développer le modèle théorique, et étudier deux domaines pertinents d’application que sont la recherche par des enzymes de séquences-cibles sur l’ADN et les trajectoires d’animaux.

3 Développements à venir du modèle de recherche intermittente

De nombreux développements théoriques sont envisagés. Je présente ici les prolongements les plus immédiats.

Une suite du travail serait l’étude de l’effet d’une répartition poissonnienne de cibles plutôt que régulière dans un espace à deux dimensions, ce qui permettrait de voir si les résultats obtenus sont robustes. Un autre prolongement serait d’étudier le modèle à trois dimensions, ce qui permettrait d’utiliser les données existantes sur les trajectoires de plancton [14].

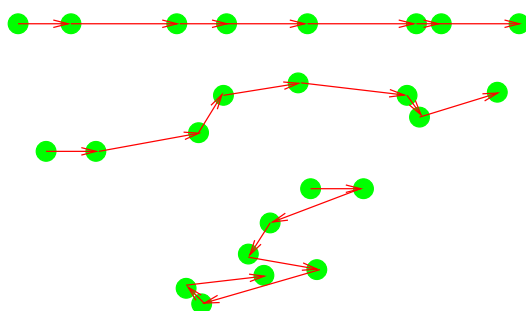


FIG. 2 – Corrélation de l’orientation des segments (en rouge) entre les pauses (en vert). En haut, la corrélation est parfaite, la trajectoire est donc une ligne droite, ce qui revient à un problème à une dimension déjà traité [7]. En bas, il n’y a aucune corrélation entre les segments, ce qui a aussi déjà été traité [11]. Au milieu, une trajectoire avec de la corrélation entre les segments, plus proche d’une trajectoire réelle de certains animaux [3].

Un autre point important est que les trajectoires des animaux sont souvent décrites comme une suite de segments d’orientations corrélées entre lesquels l’animal s’arrête pour détecter sa cible [2]. Le modèle à une dimension correspond au cas où la corrélation est très grande : on obtient une trajectoire rectiligne. Le modèle à deux dimensions sans mémoire correspond au cas où il n’y a aucune corrélation entre segments. Dans la situation réelle (cf. Fig. 2) où on a un peu de corrélation, de quel extrême se rapproche-t-on ? Sous quelles conditions ? Garde-t-on les mêmes propriétés de l’optimum des temps passés dans chaque phase ?

De nombreux autres développements peuvent être envisagés, notamment en fonctions des systèmes réels étudiés.

4 Domaine d’application : recherche par des enzymes de séquences-cibles sur l’ADN

Parmi les nombreuses protéines qui interagissent avec l’ADN, certaines sont site-spécifiques et sont capables de trouver très rapidement leur cible, souvent constituée de quelques paires de bases, sur de l’ADN comportant plusieurs dizaines de milliers de paires de base.

C’est le cas des enzymes de restriction de type II, protéines présentes dans les bactéries et destinées à lutter contre les attaques virales. Ces enzymes reconnaissent une courte séquence d’ADN du virus, et coupent le double brin en cette séquence avant que la machinerie moléculaire de la bactérie n’ait été détournée au profit du virus. Il est donc vital pour une bactérie que ses enzymes de

restriction trouvent leur séquence spécifique le plus rapidement possible. De nombreuses expériences ont été menées *in vivo* par des biochimistes pour mesurer les taux d'association de ces enzymes à leur cible. Des taux très élevés ont été relevés, qu'un modèle ne prenant en compte que la diffusion à trois dimensions des protéines ne peut expliquer. D'autre part, ces études ont révélé que ce taux d'association augmente avec la taille de l'ADN non spécifique qui borde la séquence cible. Il semble donc que ces enzymes diffusent le long de l'ADN, après s'être fixée sur de l'ADN non spécifique, ce mouvement à une dimension alternant avec des phases de diffusion à trois dimensions [1].

L'équipe qui est mon premier choix dans cette demande de bourse a déjà travaillé sur ce thème [6], et débute une collaboration avec P. Desbiolles, de l'équipe "Optique et biologie" du Laboratoire Kastler Brossel (ENS Paris). Il a élaboré avec ses doctorants A. Crut et I. Bonnet des techniques de pointe pour l'observation de molécules individuelles d'ADN et de leur interaction avec EcoRV, une enzyme de restriction de type II. [16]. Dans ces expériences, l'ADN est attaché spécifiquement par ses extrémités à une surface, sans être surtiré. Les enzymes sont liées à des fluorophores organiques ou inorganiques (nanocristaux semi-conducteurs) qui permettent de suivre leur trajectoire par microscopie de fluorescence. La diffusion de l'enzyme le long de l'ADN a été récemment mise en évidence et caractérisée par cette équipe. De nombreux "sauts" de l'enzyme le long de l'ADN ont été détectés, si bien qu'un modèle de recherche intermittente pourrait être particulièrement adapté à l'interprétation de ces données.

Un projet en cours est d'utiliser des pinces optiques pour "tenir" les extrémités de l'ADN (plus exactement des billes diélectriques attachées aux extrémités). Ce dispositif permet de résoudre le problème d'interactions non spécifique de l'ADN et des protéines avec la surface, et de pouvoir plus facilement varier l'étirement de l'ADN. Des expériences complémentaires avec un ADN contenant une séquence-cible pour EcoRV pourraient donner aussi des informations intéressantes sur la répartition du temps passé entre les différentes phases.

Si j'obtiens mon premier choix de bourse de thèse, il est prévu que je passe une partie de mon temps dans l'équipe du LKB afin de d'apprendre les techniques expérimentales (micromanipulation d'ADN, microscopie de fluorescence et détection d'objets nanométriques individuels) et que je mène des expériences dans cette équipe.

5 Autre domaine d'application : trajectoires d'animaux

Un autre domaine important d'application du modèle de recherche intermittente est l'explication de trajectoires d'animaux.

L'équipe où je fais ma demande en premier choix a commencé il y a deux ans une collaboration avec Vincent Fourcassié du Centre de Recherche sur la Cognition Animale à l'université Toulouse III. Il travaille sur les insectes, et en particulier les fourmis. Certaines, lorsqu'elles partent à la découverte d'un espace inexploré, vierge de trace olfactives laissées par d'autres fourmis, présentent une trajectoire avec des réorientations. L'étude de ces trajectoires pourrait permettre de valider le modèle de recherche intermittente. En particulier, comme le temps optimal prévu τ_{opt} à passer dans la phase rapide non-réactive semble être une propriété robuste du modèle, l'analyse se concentrerait sur la longueur moyenne des déplacements entre les réorientations $L = v\tau_{opt}$ (où v est la vitesse de déplacement dans la phase rapide). Le modèle à deux dimensions prédit que $L = \frac{a}{\sqrt{2}} \sqrt{-(1 + \ln d)}$, où a est le rayon dans lequel la cible peut-être détecté, et d la densité de cibles. a peut être mesuré [3]. d peut facilement être varié.

Cet expérimentateur est en collaboration étroite avec Richard Fournier du laboratoire d'énergétique de l'université de Toulouse III [17]. Richard Fournier est mon deuxième choix comme directeur de thèse. Ses thématiques de travail proches me permettraient aussi de mener à bien ce projet de thèse.

Références

- [1] D.M. GOWERS, G.G. WILSON et E. HALFORD : Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. *Proceedings of the National Academy of Sciences*, 102(44):15883–15888, 2005.
- [2] W.J. BELL : *Searching behaviour : the behavioural ecology of finding resources*. Chapman and Hall, 1991.
- [3] W.J. O'BRIEN, H.I. BROWMAN et B.I. EVANS : Search strategies of foraging animals. *American Scientist*, 78:152–160, 1990.
- [4] G.M. VISWANATHAN, S.V. BULDYREV, S.HAVLIN, M.G.E. Da LUZ, E.P. RAPOSO et H.E. STANLEY : Optimizing the success of random searches. *Nature*, 401(6756):911–914, 1999.
- [5] E.P. RAPOSO, S.V. BULDYREV, M.G.E. Da LUZ, M.C. SANTOS, H.E. STANLEY et G.M. VISWANATHAN : Dynamical robustness of levy search strategies. *Physical review letters*, 91(24), 2003.
- [6] M. COPPEY, O. BÉNICHOU, R. VOITURIEZ et M. MOREAU : Kinetics of target site localization of a protein on DNA. *Biophysical journal*, 87:1640–1649, 2004.
- [7] O. BÉNICHOU, M. COPPEY, M. MOREAU, P-H. SUET et R. VOITURIEZ : Optimal search strategies for hidden targets. *Physical Review Letters*, 94, 2005.
- [8] O. BÉNICHOU, M. COPPEY, M. MOREAU, P-H. SUET et R. VOITURIEZ : A stochastic model for intermittent search strategies. *Journal of physics : condensed matter*, 17:1–12, 2005.
- [9] O. BÉNICHOU, M. COPPEY, M. MOREAU, P-H. SUET et R. VOITURIEZ : A stochastic theory for the intermittent behaviour of foraging animals. *Physica A*, 356(1):151–156, 2005.
- [10] M.MOREAU, O.BÉNICHOU, C.LOVERDO et R.VOITURIEZ : Intermittent search process in disordered medium. *Soumis à Europhysics Letters*.
- [11] O.BÉNICHOU, C.LOVERDO, M.MOREAU et R.VOITURIEZ : Two-dimensional intermittent search processes : An alternative to lévy flight strategies. *Physical Review E*, 74:020102, 2006.
- [12] O.BÉNICHOU, C.LOVERDO, M.MOREAU et R.VOITURIEZ : A minimal model of intermittent search in dimension two. *A paraître dans Journal of physics : condensed matter*.
- [13] D. BOYER, G. RAMOS-FERNÁNDEZ, O. MIRAMONTES, J.L. MATEOS, G. COCHO, H. LARRALDE, H. RAMOS et F. ROJAS : Scale-free foraging by primates emerges from their interaction with a complex environment. *Proceedings of the Royal Society B*, 273:1743 – 1750, 2006.
- [14] F. BARTUMEUS, F. PETERS, S.PUEYO, C. MARRASE et J. CATALAN : Helical levy walks : Adjusting searching statistics to resource availability in microzooplankton. *Proceedings of the National Academy of Sciences*, 100(22):12771–12775, 2003.
- [15] M.F. SHLESINGER : Mathematical physics : Search research. *Nature*, 443:281, 2006.
- [16] A.CRUT : *Diffusion de l'enzyme de restriction EcoRV sur des molécules d'ADN étirées*. Thèse de doctorat, 2005.
- [17] M.CHALLET, V.FOUCASSIÉ, S.BLANCO, R.FOURNIER, G.THERAULAZ et C.JOST : A new test of random walks in heterogeneous environments. *Naturwissenschaften*, 92:367–370, 2005.