



HAL
open science

Modéliser les circuits cérébraux de la décision pour lutter contre le changement climatique

Frédéric Alexandre

► **To cite this version:**

Frédéric Alexandre. Modéliser les circuits cérébraux de la décision pour lutter contre le changement climatique. 2021. hal-03476211

HAL Id: hal-03476211

<https://inria.hal.science/hal-03476211v1>

Submitted on 12 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modéliser les circuits cérébraux de la décision pour lutter contre le changement climatique

Par Frédéric Alexandre

Pour évoquer des systèmes d'Intelligence Artificielle, on parle souvent de systèmes d'aide à la décision. Est-ce que ça veut dire que l'on sait modéliser la décision et faire des machines qui décident comme des humains ? C'est en tout cas un sujet qui intéresse l'équipe-projet Mnemosyne*, spécialisée dans la modélisation des fonctions cognitives supérieures par l'étude du cerveau. Avec des chercheurs en sciences sociales et en marketing, nous travaillons actuellement sur la prise de décision dans le contexte du changement climatique. Avant de décrire plus précisément ce projet, nous présentons les principes et les difficultés de la modélisation de la prise de décision.

Commençons par définir le processus de décision en prenant l'exemple du choix de son mode de déplacement. Choisir entre la marche à pied ou la voiture, c'est recueillir des indices en faveur ou contre chaque proposition (« je veux moins polluer » ou « on annonce de la pluie et je ne veux pas être mouillé»), observer que l'accumulation de ces indices fait pencher la balance d'un côté ou de l'autre et, à un certain moment, trancher. C'est à dire penser qu'on a une vue assez représentative de la situation pour transformer cette oscillation entre deux avis en une décision catégorique : je marche ou je prends ma voiture. Cet énoncé suggère plusieurs types de difficultés liées à la décision. Commençons par les plus évidentes. Premièrement, il faut savoir évaluer chacun des arguments présentés, les mettre sur une échelle commune et savoir les comparer (en quoi un argument compense ou me convainc plus qu'un autre). Deuxièmement, il faut avoir entendu de façon équitable et assez longtemps les deux propositions pour se faire un avis non biaisé, mais aussi, à la fin, il faut décider ; on ne peut pas rester tout le temps dans l'indécision. Il y a à trouver ici un compromis entre la vitesse et la justesse de la décision.

Des modèles mathématiques pour analyser la décision humaine

Pour ces deux types de difficultés (savoir comparer les indices et savoir trancher), les mathématiques développent des outils intéressants pour d'une part coder et comparer l'information de la façon la plus objective possible et d'autre part évaluer, pour un niveau de précision souhaité, le seuil optimal de différence entre les avis « Pour » et « Contre » qu'il faut atteindre avant de décider. Ces modèles mathématiques ont d'abord été mis en œuvre pour reproduire des tâches de décision perceptive élémentaires : on montre à un sujet des points en mouvement sur un écran et il doit décider si, globalement, ces points vont plutôt à droite ou à gauche. Il est possible de rendre cette tâche très difficile en programmant le mouvement de chacun des points avec des fonctions en grande partie aléatoires. Or on peut montrer que ces modèles mathématiques d'accumulation jusqu'à un seuil parviennent à reproduire très fidèlement la décision humaine, aussi bien dans sa performance que dans ses temps de réaction. Il reste à voir si on peut les adapter à des décisions encore plus complexes, comme le choix d'un comportement écoresponsable à partir d'indices plus subjectifs.

Des modèles neuronaux pour reproduire les circuits de la décision

Ces modèles mathématiques sont également intéressants car ils permettent de définir les grandeurs et les phénomènes critiques qui concourent à la décision comme la détection de

chaque indice, leur accumulation, la différence entre les solutions alternatives ou encore leur comparaison au seuil de décision. Des études d'imagerie cérébrale permettent d'identifier les différentes régions du cerveau impliquées dans l'évaluation de chacun de ces critères et, au cours de la décision, l'ordre dans lequel ces régions sont activées. Notre équipe de recherche travaille ensuite dans la programmation de réseaux de neurones artificiels qui, d'une part, calculent de façon similaire à ces modèles mathématiques et d'autre part, sont organisés selon une architecture globale reproduisant la circuiterie observée par imagerie et reproduisant également la dynamique d'évaluation. Nous pouvons ainsi corriger certaines limitations des modèles mathématiques évoqués plus haut. Ces derniers considèrent uniquement des décisions binaires (droite/gauche) alors que les calculs neuronaux permettent de pouvoir considérer plusieurs catégories d'indices et de réponses. D'autre part, alors que nous avons parlé jusqu'à présent de décisions mathématiquement fondées (certains disent logiques ou rationnelles), il est connu que les humains sont soumis à différents types de biais quand ils font des jugements, ce qui laisse souvent penser que nous ne sommes pas rationnels.

Les SHS au secours des mathématiques pour comprendre nos biais

En collaboration avec les sciences humaines et sociales (SHS) qui étudient et décrivent ces biais, nous cherchons à montrer que nous pouvons les expliquer et les reproduire en manipulant certains paramètres de nos modèles. Nous prétendons aussi que ces biais ne montrent pas une faiblesse de notre jugement mais plutôt une orientation de ce jugement tout à fait pertinente pour un être vivant évoluant dans des conditions écologiques.

Donnons quelques exemples : au lieu d'intégrer tous les arguments de la même manière dans notre jugement, nous pouvons être soumis à un biais de récence ou de primauté, selon que les indices les plus récents ou les plus anciens vont jouer un rôle plus important dans la décision. Ce type de jugement, qu'on peut reproduire en modifiant certains paramètres internes du calcul neuronal, peut paraître plus adapté pour certaines situations avec des conditions changeantes (récence) ou stables (primauté). On sait par ailleurs que détecter ce type de conditions nous fait émettre des neurohormones qui modifient la nature du calcul neuronal en modifiant certains de leurs paramètres.

Un autre biais, appelé aversion au risque, fait que nous surestimons les indices défavorables par rapport aux indices favorables. Il a été montré expérimentalement en économie que les neurones codent différemment les pertes et les gains, ce que l'on sait simplement reproduire dans nos modèles. Ici aussi, on peut comprendre que, pour un être vivant qui ne fait pas que calculer des bilans financiers, il est judicieux d'accorder plus d'importance à ce qui peut nous nuire (et potentiellement nous blesser) qu'à ce qui est positif (et qu'on aura d'autres occasions de retrouver). De façon similaire, des biais, appelés de référence ou de préférence, font que nous allons accorder plus d'importance à des indices faisant référence à un événement récent ou à une préférence personnelle. Ici aussi, introduire des mécanismes cérébraux connus de types mnésiques permet de reproduire ces phénomènes, dont l'intérêt adaptatif semble aussi clair. Enfin un type de biais concerne la différence entre les valeurs que nous donnons à des situations et les valeurs que nous donnons aux actions pour les atteindre ou les éviter et que l'on connaît bien en addictologie : alors que l'on sait très bien que alcool, tabac et autres drogues sont mauvais pour nous, diminuer nos actions de consommation n'est pas forcément simple. On sait que dans le cerveau ces valeurs sont dissociées.

Une collaboration régionale pour mieux comprendre les décisions humaines liées au changement climatique

Le projet que nous avons mentionné, soutenu par la Région Nouvelle Aquitaine, Inria et les universités de Bordeaux et de La Rochelle, vise à mieux comprendre les décisions humaines relatives au changement climatique et à la transition écologique afin d'aider à changer nos comportements. On peut constater que la plupart des biais mentionnés plus haut s'appliquent : est-on prêts à changer d'habitude maintenant (prendre le bus plutôt que la voiture) pour des résultats à long terme (modifier la température moyenne dans 30 ans), quitter des comportements faciles et observés autour de soi (prendre sa voiture comme tout le monde) et mettre en œuvre nos convictions citoyennes (différence entre ce qu'on pense et ce qu'on fait), etc. Avec nos collègues de sciences humaines et sociales, nous avons orienté notre approche selon les convictions suivantes : d'une part, si nous savons bien décrire les caractéristiques de cette prise de décision et en particulier pourquoi et quand elle est difficile, nous pouvons proposer des mises en situations qui seront plus favorables à des prises de décision responsables. D'autre part, nous faisons également le pari du citoyen informé et proactif. Si nous décrivons ces mécanismes de décision, leurs biais et leur inscription cérébrale, cela peut permettre de déculpabiliser (ce sont des mécanismes biologiques à la base de ces biais) et de donner des leviers pour travailler sur nos processus mentaux et réviser nos modes de pensée.

Des éléments de ce projet régional ont été présentés lors du premier Climackaton de l'université de Bordeaux :

<https://www.u-bordeaux.fr/Actualites/De-la-vie-de-campus/Premier-climackathon-pour-l-universite-de-Bordeaux>

* Mnemosyne est une équipe-projet Inria au sein du LaBRI (UMR 5800, CNRS - université de Bordeaux - Bordeaux INP) et de l'IMN (UMR 5293, CNRS - université de Bordeaux)