



20 juin 2014

## Apprentissage, motivation, émotion : comment apprenons-nous ?

Le cerveau est l'objet le plus complexe de l'univers connu, et c'est de sa compréhension que dépendra l'avenir de nos technologies futures et singulièrement l'intelligence artificielle et la robotique. A l'Inria, Frédéric Alexandre dirige le projet *Mnémosyne*, chargé de développer des modèles computationnels de notre fonctionnement cérébral, dans le but de développer à terme des agents autonomes, robotiques ou logiciels, qui soient physiologiquement crédibles. Autrement dit, dont l'architecture sera analogue à celle du cerveau humain. Au cours d'une conférence à Bordeaux, lors d'un atelier organisé conjointement par l'expédition *Bodyware* de la Fing et l'Inria, Frédéric Alexandre nous a présenté quelques fondamentaux concernant les mécanismes de notre mental.

Les sciences de l'esprit, sciences cognitives et neurosciences, ont fait des progrès ces dernières années et s'attaquent à des sujets de plus en plus difficiles, a-t-il expliqué en préambule. Ainsi, dans le domaine de la perception et du « machine learning », on est passé d'une étude de la reconnaissance des formes à celle de l'interaction avec l'utilisateur.

En neurosciences, on va de la recherche sur la vision à celle des fonctions exécutives. Enfin, dans le domaine de la cognition, de l'analyse du raisonnement à celle de la cognition incarnée (dont vous trouverez une définition [dans cet article](#)).

Mais finalement, quelle est la vraie question à se poser sur l'esprit humain ? C'est de se demander comment s'adapter dans un monde dynamique et imprévu, alors que l'on dispose de très peu d'aide pour cela. Pour répondre à cette interrogation, Frédéric Alexandre a souligné l'importance pour nos fonctions cognitives de trois processus fondamentaux interagissant entre eux : apprentissage, motivation, et émotion. Ces trois processus impliquant des relations entre le cerveau, le corps et l'environnement.

### Comment garder l'information importante

Commençons par l'apprentissage. En quoi consiste-t-il exactement ?

Il s'agit de pouvoir extraire des invariants dans un monde multimodal et multisensoriel. Par exemple, si je vois un chien aboyer ou je l'entends aboyer je sais que j'ai affaire au même phénomène, bien que les modes de perception soient différents ; je vois ma main bouger et je touche un objet. Je comprends que ces deux actions sont liées.

L'apprentissage, c'est aussi adapter une connaissance initiale ou acquise aux dérives et aux changements. Un même objet ne se présente pas toujours sous le même angle, n'a pas toujours le même comportement, voire la même apparence (par exemple, le rapport entre un yorkshire et un pitbull est assez lointain, mais ce sont tous les deux des chiens)

Nous recevons bien trop d'informations en provenance du monde extérieur. Il faut effectuer des choix. En gros, trois stratégies sont utilisées par le cerveau. Réduire les data, ce qui se fait

par la création d'un prototype : autrement dit, on élimine la plupart des exemples d'un même phénomène pour en retenir un qui est le plus représentatif de la catégorie mentale qu'il représente. Par exemple une table qui apparaît comme représentative de toutes les tables, le prototype de la table.

On peut également réduire la dimensionnalité par le processus de représentation. En effet, chaque perception possède une multitude de paramètres : la table peut être construite dans un certain matériau, posséder une certaine couleur, se trouver dans un lieu donné, porter des assiettes ou un vase. Autant de « dimensions » qu'on peut éliminer pour arriver au concept de « table ».

## **Le rôle des émotions**

Il existe une troisième stratégie, et celle-ci implique les émotions : le cerveau peut augmenter des signaux spécifiques internes ou externes, comme la douleur le plaisir, la nouveauté...

Les émotions font que tous les éléments ne sont pas équivalents : ils n'ont pas la même importance. C'est la différence existant entre une cognition « froide », rationnelle, et une cognition chaude, émotionnelle. Cette dernière fonctionne suivant un schéma très binaire : les choses sont plaisantes ou déplaisantes, bonnes ou mauvaises...

Certaines de ces émotions sont innées ; par exemple la peur du vide, ou celle d'un serpent sont probablement inscrites dans notre ADN. La crainte d'une arme à feu, s'avère, au contraire, bien évidemment apprise.

On peut définir une émotion comme une information simple qui va avoir un impact énorme sur le corps. Elle peut en fait servir à deux choses. Tout d'abord à évaluer l'intensité, l'importance d'un événement : le petit frisson que je peux éprouver parfois en regardant le Projet Blair Witch n'est pas du même ordre que la peur que je ressentirais si je me trouvais face à un cobra royal prêt à me sauter dessus. Mais l'émotion est également un signal social : je rougis, je tremble je bafouille, mon visage prend certaines expressions particulières. En signalant mon émotion à mon entourage social j'envoie très vite une information qui peut être utile à la survie de tous. Je peux signaler la présence d'un danger imminent bien plus rapidement que si je recourais au langage, par exemple.

Il existe plusieurs manières de catégoriser les émotions, l'une des plus connues, que nous a montrée Frédéric Alexandre, est la « [roue de Plutchick](#) ».

## **Des émotions aux motivations**

Il existe deux types de motivations : les motivations extrinsèques, qui sont renforcées par des demandes venant de l'extérieur de notre esprit : faim soif, sexe, intégrité du corps...

Et il y a les motivations intrinsèques, d'ordre psychologique, comme la nouveauté, l'erreur de prédiction, exploration, la curiosité...

Cette notion d' « erreur de prédiction » est intéressante, d'autant qu'à un autre moment de sa conférence, Frédéric Alexandre a associé à ce phénomène la très célèbre dopamine. Mais celle-ci n'est-elle pas liée à la récompense, au plaisir d'avoir réussi quelque chose ? Eh bien les deux aspects sont liés, explique Alexandre : lorsque nous réussissons quelque chose pour la première fois, nous éprouvons en effet une sensation de surprise (si nous n'avions pas cette réaction, notre « réussite » serait la simple exécution d'une tâche), donc le résultat ne correspond pas à la « prédiction ».

On peut être tenté de confondre motivation et émotions qui sont pourtant fondamentalement différentes. Les émotions donnent une valeur aux choses, mais elles demeurent passives. Au contraire, les motivations nous poussent à l'action : elles nous font dire : « Je veux... Je ne veux pas » et non « J'aime » ou « J'aime pas »...

Les techniques de conditionnement permettent d'agir sur les émotions et les motivations. Le premier type de conditionnement, le conditionnement pavlovien est un apprentissage émotionnel des stimuli neutres qui annoncent des récompenses et des douleurs. Par exemple on fait sonner une cloche et on envoie une décharge électrique au pauvre animal. Dans ce type de cas, on prédit, mais on subit quand même.

Dans le cas du conditionnement opérant, au contraire, on provoque une réponse de la part du sujet. Par exemple, le rat est capable de presser une manette pour obtenir de la nourriture. C'est un apprentissage motivationnel.

Lorsqu'on essaie de comprendre cette intrication d'émotions et de motivations, on se heurte à des questionnements, des interrogations : je meurs de faim, je vois un croissant mais il est situé derrière un serpent, que dois-je faire ? Comment comparer des motivations ou des émotions, et évaluer leur importance respective ? Il faut pour cela une monnaie commune, et c'est le corps qui fournit cet étalon. Selon ce que je ressens au niveau biologique (ma faim, ma peur) je peux hiérarchiser les différents ressentis.

Cette réflexion sur les émotions, les motivations et l'apprentissage n'est pas exclusivement abstraite. Elle correspond à des données physiologiques qui peuvent être situées dans le cerveau.

Au centre de tout apprentissage, il y a la mémoire ou plutôt les mémoires. Tout d'abord il faut distinguer la mémoire déclarative (lorsque je sais quelque chose et que je peux l'exprimer verbalement : je sais que le ciel est bleu) et non déclarative (je sais faire du vélo). La « mémoire de travail », elle, est une mémoire à court terme qui permet de se rappeler les événements récents. Elle est indispensable à l'exécution des motivations, puisqu'elle permet de se rappeler les buts qu'on poursuit. Elle est localisée dans le [cortex préfrontal](#). L'[hippocampe](#) stocke lui les souvenirs dans une mémoire épisodique. Autrement dit, c'est grâce à lui qu'on se souvient de ce que nous voyons et faisons. Dans le cortex postérieur se déroulent les opérations de [mémoire sémantique](#). C'est là que se situe le « concept de chaise » élaboré à partir des exemples stockés dans l'hippocampe, que ce soit par la « réduction des data » ou la « réduction dimensionnelle ». Le conditionnement pavlovien se trouve, qui s'en étonnera, dans l'[amygdale](#), qui gère les émotions, les réactions de « fight or fly » comme disent les anglo-saxons (combattre ou s'enfuir). Mais le conditionnement opérant, lui, est géré par une autre zone, les [ganglions de la base](#).

Aux zones du cerveau s'adjoignent aussi les neurotransmetteurs qui jouent leur rôle dans l'apprentissage et la gestion de la mémoire et de l'apprentissage. On a déjà vu ce que faisait la dopamine. La sérotonine, elle, permet de calculer le risque que je peux prendre. La noradrénaline est associée aux incertitudes connues, l'acétylcholine aux incertitudes inconnues...

Pour effectuer ses travaux, l'équipe Mnémosyne utilise bien sûr de véritables data physiologiques obtenues par l'armada des capteurs existants.

La recherche sur les « agents autonomes » n'est pas cantonnée à Mnémosyne au sein de l'Inria. Pierre-Yves Oudeyer, également présent lors de cette journée de travail, s'y

consacre aussi, mais en suivant un autre voie, qu'il avait déjà exposée lors de sa récente conférence à Lift Marseille . Toutefois sa démarche diffère, elle se situe à un autre niveau d'organisation. Alors que Frédéric Alexandre et son groupe poursuivent une recherche très proche des neurosciences et travaillent à élaborer le modèle le plus proche possible d'un cerveau biologique, Pierre-Yves Oudeyer s'intéresse plutôt à l'aspect psychologique et notamment développemental, dans la lignée des travaux d'un Piaget : comment l'intelligence vient aux bébés, et un jour peut-être, aux robots.

Restera aux chercheurs du futur à synthétiser en une seule théorie toutes ces approches ! Ce jour là, peut être pourrons nous parler véritablement d'intelligence artificielle et voir apparaître des machines dotées de motivations, d'émotions et capables d'un « vrai » apprentissage. Jusqu'où ces robots seront ils basés sur l'architecture biologique de l'être humain ? Et leur esprit sera-t-il, comme le nôtre, inscrit dans leur corps ?

Rémi Sussan

