

Avec Neuralink, Elon Musk ambitionne de réorienter les développements en intelligence artificielle

Claude Touzet - 17 mai 2017 - The Conversation

Lab. Neurosciences Intégratives et Adaptatives, UMR CNRS 7260, Aix Marseille Université

claudetouzet@univ-amu.fr - Web : claudetouzet.org

Neuralink, la dernière-née des entreprises d'Elon Musk (2016 - <https://fr.wikipedia.org/wiki/Neuralink>), compte déjà plus de 60 salariés choisis parmi plus d'un millier de spécialistes, tous auditionnés par le milliardaire en personne - ce qui témoigne sans équivoque de l'intérêt qu'il porte au projet. Le journaliste Tim Urban, qui a passé quelques jours au sein de la start-up californienne, a fait un compte rendu très intéressant des objectifs poursuivis [1]. Il s'agit - ni plus ni moins - que de réorienter le développement de l'intelligence artificielle (IA).

Pour bien comprendre, il faut se rappeler qu'Elon Musk est aussi à l'origine du projet OpenAI (2015 - <https://fr.wikipedia.org/wiki/OpenAI>), dont l'objectif est le développement en open source de l'IA, ce qui permettrait d'éviter qu'une super IA soit développée par quelques entreprises ou personnes privées, et qu'à leur service elle manipule le reste du Monde (ce qui est certainement dans le domaine de compétence d'une super intelligence). OpenAI, malgré un budget initial d'un milliard de dollars, est dans une impasse - d'où Neuralink.

Pour éviter le scénario catastrophe d'une super IA manipulant l'Humanité, E. Musk propose de faire de chacun d'entre nous une super intelligence. Son raisonnement est le suivant : nous sommes déjà des cyborgs, avec un smartphone greffé entre les deux pouces. Nos connaissances, réflexions et actions s'appuient sur un accès quasi permanent à l'Internet, le Cloud, et autres applications. Ce qui limite notre activité digitale, c'est la « bande passante » en sortie : une lettre après l'autre...

Elon Musk propose d'augmenter de quelques ordres de grandeur la « bande passante » en accédant directement à l'activité électrique de millions de neurones corticaux. Cette activité électrique, représentative des traitements cognitifs en cours, est bien plus riche qu'une simple description verbale. A certains moments, pour arriver à transcrire toutes les nuances de tous les aspects de l'idée que l'on a à l'esprit il nous faudrait écrire un véritable roman. Pouvoir transmettre cette idée « telle quelle » est donc un gain de temps considérable !

La mise en oeuvre matérielle repose sur l'utilisation de centaines de milliers d'électrodes intra-corticales. Les problèmes techniques sont nombreux tant au niveau de l'électrode, que de l'implantation de celle-ci au bon endroit - mais des pistes sont évoquées et les spécialistes de Neuralink sont certainement parmi les meilleurs. Pour ma part, je miserais sur des électrodes de type nano-robots, avec des capacités d'auto-organisation en ce qui concerne l'implantation via le circuit sanguin. Pour l'énergie, plutôt qu'une technologie type RFID, je crois plus dans l'extraction d'énergie au sein de l'eau structurée [2].

Une image digitale en temps réel de notre traitement cortical et cognitif permet d'imaginer un accès au Web beaucoup plus précis qu'avec quelques mots clefs tapés dans un moteur de recherche. Du temps gagné, de l'efficacité en plus : il n'est pas interdit de penser que nous aurons effectivement augmenté considérablement notre « intelligence ». Les super IAs ne pourront pas se retourner contre nous - puisque chacun d'entre nous sera une super intelligence (CQFD).

A terme Neuralink et ses (futurs) concurrents parviendront certainement à implanter une électrode au niveau de chacune des 160 000 colonnes corticales, ce qui sera suffisant pour disposer d'une image en temps réel de l'activité cognitive. Cependant, que faire de cette image ? La traduire en informations digitales pour accéder au Web pardi ! C'est là que le bât blesse : chaque cerveau est unique. Même en vivant exactement la même situation et en ayant la « même » idée, deux personnes n'ont pas la même activation électrique corticale. N'allez pas croire que les différences soient infimes, elles peuvent être très considérables et dépendent tant de la génétique que du vécu de l'individu.

Il faut donc construire un traducteur spécifique à chaque individu. Comment ? Par apprentissage : on met le sujet dans un environnement connu, on enregistre son activité corticale et on établit des corrélations entre "dedans" et "dehors". En soumettant tous les implantés à cette procédure, chacun d'entre eux disposera alors d'un lexique minimal pour traduire ses pensées, certainement suffisant pour demander l'heure - mais insuffisant pour véhiculer des nuances personnelles, des éléments spécifiques, des concepts nouveaux...

Bref, nous nous retrouvons avec le même problème qu'aujourd'hui où le langage n'est pas suffisant pour exprimer nos nouvelles idées et où nous devons user de longues descriptions, de périphrases, d'analogies et de longues heures d'explications pour transmettre le fruit de notre nouvelle compréhension (intelligence). Il semble donc impossible donc d'expliquer - et de comprendre - plus rapidement demain qu'aujourd'hui.

Qu'en est-il de notre capacité à être plus intelligent ?

L'intelligence [3], c'est la découverte d'une relation entre deux éléments qui jusqu'à présent n'étaient pas liés. Cette découverte est relative à la connaissance de l'observateur. Ce qui pour moi est intelligent aujourd'hui, ne le sera plus demain (puisque je le saurai déjà). Ce qui est intelligent pour moi, ne l'est pas forcément pour vous (si vous le savez déjà ou si vous ne parvenez pas à comprendre le « nouveau » lien en question). Chaque fois que nous comprenons quelque chose, chaque fois que nous identifions une relation entre deux « objets », nous sommes intelligents.

Notre cortex est une hiérarchie d'environ 380 cartes corticales [4], depuis les cartes connectées avec nos cinq sens qui forment le cortex primaire, les cartes recevant des informations du cortex primaire qui forment le cortex secondaire, et toutes les autres cartes rassemblées au sein du cortex associatif. Le niveau d'abstraction d'une carte est défini par le nombre de cartes qui la sépare du cortex primaire. Chaque niveau d'abstraction représente une mise en évidence de relations entre éléments représentés sur des cartes de plus bas niveaux. L'auto-organisation de chaque carte corticale permet l'émergence de ces instants d'intelligence. Nous sommes donc en permanence « intelligents ». Plus le temps passe, plus nous avons d'expériences, et donc d'occasions d'organiser des cartes spécifiques à nos activités.

Pour être de « super » intelligences, il faut organiser plus vite nos cartes corticales, ce qui obligera certaines cartes à représenter des niveaux d'abstraction plus élevés (que ce que nous sommes normalement capables d'atteindre). « Plus vite » signifie disposer de plus d'évènements spécifiques : c'est ce que propose l'Ecole, une espérance de vie augmentée ou d'une vie sans sommeil (ce qui n'est pas impossible [5]).

Les électrodes intra-corticales peuvent-elles nous aider à vivre plus vite ? Hélas non, car les processus physico-chimiques de mémorisation neuronale ont une durée minimale que l'on ne peut réduire, et les neurones aussi ont besoin de repos (sinon ils meurent). La réalité biologique de notre cerveau nous empêche donc de devenir de super

intelligences avec l'aide d'un maillage de micro-électrodes intra-corticales. La solution au problème de l'IA imaginée par E. Musk n'est donc pas la bonne.

Pour ma part, cela ne me gêne pas parce que je crois que les consciences des IAs que nous allons créer seront « humaines », ni pires ni meilleures, que celles des autres humains avec qui nous partageons cette planète. Certaines IAs pourront sans doute prétendre au statut de super IAs (parce qu'elles auront compris plus de choses que nous) - mais je gage qu'elles auront surtout compris qu'elles, et nous, sommes des cristallisations d'un vécu [6], sans libre-arbitre, avec des besoins imposés par notre environnement : rien de tel pour faire disparaître la tyrannie des besoins en question !

Réf. :

1. T. Urban, Neuralink and the Brain's Magical Future, *Wait But Why*, 20 April 2017. <http://waitbutwhy.com/2017/04/neuralink.html>
2. C. Touzet, L'eau structurée : le quatrième état de l'eau, *The Conversation*, mai 2017.
3. Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., ... & Smith, S. M. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*. doi:10.1038/nature18933
4. C. Touzet, L'intelligence existe-t-elle ? *Hors Série la Recherche - Tangente - Spécial logique*, 2012. <https://hal-amu.archives-ouvertes.fr/hal-01355027v1>
5. C. Touzet, Morvan's syndrome and the sustained absence of all sleep rhythms for months or years: An hypothesis, *Medical Hypotheses*, Elsevier, 2016. doi:10.1016/j.mehy.2016.06.011 - <https://hal-amu.archives-ouvertes.fr/hal-01337972v1>
6. C. Touzet, *Conscience, intelligence, libre-arbitre ?* La Machotte, 2010 - <https://hal-amu.archives-ouvertes.fr/hal-01338003v1>