

Les techniques instrumentées, du calcul formel à l'IA générative

Jean-baptiste Lagrange

LDAR, France ; lagrange@math.univ-paris-diderot.fr

Résumé. L'essor de l'IA générative entraîne une mutation des pratiques d'enseignement apprentissage des mathématiques comparable à celle qui s'est produite lors de l'apparition du calcul formel dans les années 1990. Un rapprochement est établi en s'appuyant sur des situations didactiques impliquant la factorisation des polynômes. Dans ce texte à dominante théorique l'accent est mis sur les techniques instrumentées et leur articulation avec la théorie des espaces de travail mathématiques.

Les techniques instrumentées sont étudiées sur des situations précises – la dialectique construction molle construction robuste avec la Géométrie Dynamique, et la factorisation des polynômes avec le calcul formel – s'appuyant sur une analyse en profondeur du fonctionnement de ces artefacts dans le domaine auquel la situation se rapporte. L'analyse s'appuie sur les idées de pseudo-transparence des outils numériques et de valences pragmatique et épistémique des techniques, ainsi que de leur viabilité institutionnelle. Le même type d'analyse est mené sur la factorisation avec l'IA générative. Il conduit à des pistes pour la mise en place de situations didactiques.

Pour finir, le texte s'ouvre sur des enjeux plus larges pour l'enseignement des mathématiques, notamment les dimensions mathématique, écologique et sociale de l'IA générative.

Mots clés : Techniques instrumentées, calcul formel, IA générative, factorisation des polynômes, valences pragmatique et épistémique, viabilité institutionnelle, espaces de travail mathématique.

Abstract. The rise of generative AI is bringing about a shift in math teaching practices comparable as when computer algebra came out in the 1990s. A link is made based on teaching situations involving polynomial factorization. This mainly theoretical article highlights the idea of instrumented techniques and how they fit in with the theory of mathematical workspaces.

Instrumented techniques are studied in specific situations: the dialectic between soft construction and hard construction with Dynamic Geometry, and the factorization of polynomials with computer algebra, based on an in-depth analysis of how these artifacts function in the domain to which the situation pertains. The analysis is based on the ideas of pseudo-transparency of digital tools and the pragmatic and epistemic valences of techniques, as well as their institutional viability. The same type of analysis is carried out on factorization with generative AI. It leads to suggestions for the implementation of didactic situations.

Finally, the text opens up broader issues for mathematics education, in particular the mathematical, ecological, and social dimensions of generative AI.

Keywords: Instrumented techniques, computer algebra, generative AI, polynomial factorization, pragmatic and epistemic valences, institutional viability, mathematical working spaces.

Introduction

Ce texte part de situations didactiques étudiées dans les années 1990 au moment où le calcul formel devenait disponible dans les classes sous forme d'un logiciel puis de calculatrices dites symboliques. Je cherche à établir un lien avec l'intelligence artificielle générative telle qu'elle est aujourd'hui disponible pour les élèves, de façon à éclairer les analyses et débats sur cette innovation et ses conséquences dans l'enseignement des mathématiques.

Pourquoi le calcul formel et pourquoi ces situations ? Dans les années 1990, l'introduction du calcul formel dans l'enseignement des mathématiques a constitué une rupture analogue à celle constatée aujourd'hui dans beaucoup de disciplines avec l'IA générative. Plus encore que les calculatrices 4 opérations et les logiciels de géométrie, le calcul formel était vu comme capable d'effectuer des tâches comme la factorisation ou le calcul de dérivées ou d'intégrales, dévolues jusque-là aux élèves et que ceux-ci ne pouvaient accomplir qu'au prix d'un apprentissage laborieux. Ces tâches étaient souvent considérées dans la littérature sur l'enseignement des mathématiques comme des points obligés pour une pratique mathématique, mais pas nécessairement significatives en soi. Il était donc proposé par des chercheurs et des innovateurs de mettre à profit le calcul formel pour libérer les élèves de tâches jugées non significatives et permettre ainsi à ceux-ci d'accéder directement à des tâches considérées comme importantes pour les conceptualisations mathématiques. Parallèlement beaucoup de professeurs et même d'élèves étaient loin d'être enthousiastes à l'idée d'un tel changement qui leur semblait intuitivement remettre en cause des fondamentaux de l'enseignement.

Les situations sur la factorisation des polynômes construites et menées par Mounier et Aldon (1996) ont fourni une base pour une réflexion didactique et ont conduit à un cadre théorique centré sur les techniques instrumentées, inspiré de la théorie anthropologique du didactique.

Pourquoi l'intelligence artificielle générative ? Il existe de multiples formes d'intelligence artificielle. Parmi celles qui concernent les mathématiques, le calcul formel occupe une place de premier plan, mais il existe d'autres applications qualifiées de systèmes experts qui font partie comme le calcul formel de l'IA symbolique. L'IA générative relève quant à elle de l'apprentissage machine basé notamment sur les réseaux neuronaux et le renforcement. Lagrange et al. (2023) font le point dans un chapitre de handbook sur différents systèmes d'IA principalement produits par la recherche en éducation mathématique. Ce texte propose de s'intéresser plutôt à ce qui est produit par l'industrie sans intention directe quant à une utilisation dans l'enseignement, c'est-à-dire les grands systèmes d'IA générative (LLM en anglais), en faisant l'hypothèse que ce sont ces systèmes que les élèves vont rencontrer plutôt que des produits de la recherche. Par ailleurs, des systèmes directement destinés à l'enseignement, mis en place par des institutions éducatives¹ ou par des sociétés commerciales ou même des systèmes issus de la recherche vont utiliser les capacités de l'IA générative en intégrant des agents conversationnels issus des LLM. Ainsi, des caractéristiques des LLM vont nécessairement se retrouver dans ces systèmes.

¹ Par exemple MIA Seconde.

L'irruption de l'IA générative dans l'enseignement des maths constitue un phénomène analogue à l'apparition du calcul formel dans les années 1990, mais à une bien plus grande échelle. Selon la revue en ligne legrandcontinent.eu, l'usage des LLM diminue significativement pendant les congés scolaires², ce que la revue considère comme une indication de très larges usages par les élèves et les étudiants. Il existe aussi une forte perplexité des institutions scolaires et universitaires face à ces usages³. En revenant trente ans en arrière, ceci remet en mémoire l'observation par Trouche (1994) d'un hiatus entre des usages débridés chez les élèves et une difficulté du côté de l'institution à prendre en compte ces usages à propos des calculatrices.

Pour les élèves, une calculatrice graphique (dé)montre les propriétés des fonctions. Pour les professeurs une calculatrice graphique est un gadget qui n'a aucune légitimité mathématique. Les uns et les autres s'accordent cependant sur un fait : la maîtrise de cet outil ne nécessite aucun apprentissage particulier. (Trouche, 1994, page 39)

Par ailleurs, dans la recherche en éducation mathématique, Lagrange et al (2003) ont observé un foisonnement de publications centrées sur les potentialités du calcul formel et des calculatrices.

Avec les LLM aujourd'hui, un même hiatus est observé entre les usages par les élèves et une difficulté du côté de l'institution. Dans la recherche en éducation mathématique la publication d'articles se développe aussi (Pepin et al 2025). Les articles restent centrés sur des potentialités de ces systèmes à un niveau général, même si les auteurs montrent davantage de prudence quant à la possibilité de voir ces potentialités se réaliser.

L'expérience de 30 ans de recherche sur les usages des technologies numériques me conduit à penser qu'un travail de terrain au plus près des pratiques est indispensable. Mais il doit s'accompagner d'un travail théorique visant à élucider les potentialités des technologies et les conditions de leur actualisation dans la diversité de leurs applications. Ce texte vise à y contribuer.

Je rappelle d'abord dans une première section les situations sur la factorisation des polynômes de Mounier et Aldon (1996) en les situant dans le cadre des espaces de travail mathématiques. Je rapporte ensuite dans une seconde section une expérience d'usage personnel de deux LLM dans le même domaine, de façon à initier une étude du fonctionnement de ces systèmes par comparaison avec le calcul formel. Ces deux sections mettent en évidence des techniques d'usage spécifiques tant du calcul formel que des LLM et la section suivante précise la genèse de la notion de technique instrumentée et comment elle peut s'articuler avec la théorie des

² En juin et juillet 2025, les utilisateurs de Chatgpt génèrent en moyenne 23 milliards de token. Le 2 septembre 2025, la plateforme a enregistré un pic à plus de 65 milliards de token. <https://legrandcontinent.eu/fr/2025/09/03/la-croissance-de-lia-est-elle-principalement-soutenue-par-les-etudiants-et-les-adolescents/> De plus la France est un des pays où l'IA générative est la plus utilisée. <https://legrandcontinent.eu/fr/2026/01/13/les-etats-unis-et-la-chine-sont-en-retard-sur-leurope-en-matiere-dadoption-de-lia/>.

³ Voir la section « Des usages qui soulèvent des interrogations » dans un récent document ministériel « Cadre d'usage de l'IA en éducation » <https://www.education.gouv.fr/media/227697/download>.

espaces de travail mathématiques. Ensuite une section approfondit l'analyse du calcul formel et des LLM de façon à montrer comment la nécessité de techniques instrumentées trouve son origine dans le fonctionnement de ces systèmes.

Situations d'usage du calcul formel à propos de la factorisation des polynômes

L'objectif général des situations était pour des élèves de 2^{nde} d'acquérir des connaissances relatives à la factorisation des polynômes à une variable. Les connaissances sont d'ordre sémiotique (vocabulaire, conventions d'écriture...), d'ordre discursif (lien avec les racines, factorisations remarquables...) et instrumentales (techniques de factorisation). Le travail concret a porté sur un domaine typique : les polynômes de la forme $x^n - 1$. Il y a eu trois versions, la première en papier crayon et les deux autres avec un logiciel de calcul formel (DERIVE). Dans les trois versions, la tâche consistait à effectuer des factorisations de polynômes de la forme $x^n - 1$ pour diverses valeurs de n , de façon à émettre des conjectures sur des factorisations générales.

Dans une première version sur une heure en papier crayon, les élèves ont trouvé le facteur $x - 1$, puis par division sur quelques exemples conjecturé le second facteur ($x^{n-1} + \dots + x + 1$) et ainsi obtenu une factorisation en deux facteurs que je qualifie par commodité de *générique* dans la suite. La technique de division a été laborieuse et peu d'élèves ont cherché d'autres conjectures. La tâche a été surtout dans la dimension instrumentale, mais la technique mise en œuvre n'a pas permis le passage à la dimension discursive. C'est ce qui a motivé les professeurs pour une utilisation du calcul formel dans une seconde version de la tâche.

Cette seconde version a mobilisé surtout la visualisation, la genèse instrumentale étant limitée par le caractère presse-bouton de la technique d'obtention des factorisations via la commande *factor* de DERIVE. En s'en tenant aux factorisations complètes (i.e. en facteurs irréductibles) obtenues de ce logiciel de calcul formel, les élèves ont été conduits à remettre en cause la généralité de la factorisation générique pour la réserver au cas n premier et n'ont pas pu dégager d'autres conjectures, les cas n non premier ne présentant pas de régularité pour les valeurs observées (4, 6, 8, 12...). L'obstacle cognitif est une unicité supposée de la factorisation.

Cet obstacle s'analyse grâce à la notion de *pseudo transparence* introduite par Artigue (1997) pour rendre compte des premières observations d'élèves utilisant le calcul formel. Le calcul formel utilise en effet des représentations internes et des algorithmes non directement accessibles ou compréhensibles par l'utilisateur. Pour un mathématicien ceci ne constitue pas un obstacle : ses connaissances lui permettent de trouver les bonnes commandes et d'interpréter les résultats. Dans l'exemple de la factorisation des polynômes, le mathématicien comprend que le système rend une factorisation en facteurs irréductibles du fait de contraintes de représentation interne (forme standard unique qui permet de considérer plusieurs formes, par exemple des ordonnancements différents des facteurs, comme équivalentes) et externe (consistance avec l'ergonomie générale), mais qu'il existe aussi des commandes qui permettent d'en déduire des factorisations partielles. C'est précisément cette connaissance qui manque aux élèves et qui les conduit à considérer que la factorisation générique n'est vraie

que pour n premier. Ce qui est transparent pour le mathématicien (une factorisation complète en p facteurs irréductibles porte en germe $B_p - 2$ factorisations partielles, B_p étant le nombre de partitions d'un ensemble à p éléments distincts) ne l'est pas pour l'élève.

Pour franchir l'obstacle, les professeurs ont choisi dans une troisième version de mettre l'accent sur l'obtention de factorisations partielles à partir de factorisations complètes, en sélectionnant plusieurs facteurs dans ces factorisations complètes et en les développant pour obtenir un nouveau diviseur du polynôme initial. Par exemple la factorisation générique de $x^n - 1$ pour n non premier est obtenue en développant la partie de la factorisation complète excluant le facteur $(x - 1)$. Le calcul formel apporte une aide pour obtenir d'abord la factorisation complète puis la forme développée d'un produit de facteurs irréductibles. A partir de $n = 6$ (deux facteurs irréductibles du premier degré et deux facteurs irréductibles du second degré) cette aide devient indispensable, surtout si l'on prend en compte qu'il s'agit d'élèves de Seconde (grade 10). Néanmoins les manipulations nécessaires pour sélectionner des facteurs non nécessairement consécutifs dans la factorisation partielle et développer leur produit n'ont rien de trivial et les professeurs ont pris conscience de la nécessité d'intégrer ces manipulations dans la situation.

Table 1 : Productions d'élèves à l'issue de la troisième situation (Mounier et Aldon, 1996).

1	Pour tout n , on a : $x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \dots + x + 1)$ Preuve faite
2	Pour tout n puissance de 2 : $x^n - 1 = (x-1)(x+1)(x^2+1)\dots(x^{n/2}+1)$ Preuve faite
3	Pour tout n puissance de 3 : $x^n - 1 = (x-1)(x^2+x+1)(x^6+x^3+1)\dots(x^c+x^d+1)$ où $c = 2n/3$ et $d = n/3$ Pas de preuve
4	Pour tout nombre premier, on a la seule factorisation : $x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \dots + x + 1)$ Pas de preuve
5	Pour tout n pair, on a : $x^n - 1 = (x^{n/2} + 1)(x^{n/2} - 1)$ preuve faite
6	Pour tout $n = wk$, on a : $x^n - 1 = (x-1)(x^{w-1} + x^{w-2} + \dots + x + 1)(x^{n-w} + x^{n-2w} + \dots + 1)$ preuve faite.

La troisième version a donc commencé par une initiation en classe à la manipulation des facteurs dans le logiciel de calcul formel en lien avec l'idée de la pluralité des factorisations d'un même polynôme. Puis les élèves ont été invités à pratiquer hors classe et à présenter leurs conjectures en classe à des intervalles réguliers. Le tableau 1 donne une indication des conjectures émises par les élèves. Certaines conjectures (2, 3, 4) découlent d'une généralisation des résultats obtenus dans le logiciel et d'autres (1, 5 et 6) prennent en compte des factorisations partielles obtenues par la technique de regroupement des facteurs. Je

montrera dans la suite comment ce travail des élèves peut s'analyser dans une genèse instrumentale, associée aux genèses sémiotique et discursive et comment la manipulation de facteurs d'une factorisation complète et l'obtention de factorisations partielles par développement du produit de certains facteurs s'analyse comme des *techniques instrumentées* et le rôle de ces techniques dans la genèse instrumentale.

Une exploration de l'interaction avec des LLM pour la factorisation des polynômes

Des documents d'échange de pratique écrits par des enseignants experts (notamment Durand 2025) montrent qu'une réflexion avec les élèves sur le comportement de LLM standards a une certaine productivité, et je base donc mon observation sur le même principe. Ainsi, après quelques interactions avec ChatGPT et MistralAI sur le thème des factorisations de $x^n - 1$, je tente de dégager certains invariants :

La pseudo-transparence, une caractéristique commune avec le calcul formel

Le terme de transparence situe le logiciel comme une fenêtre à travers laquelle l'utilisateur tente de cerner un phénomène tout en superposant à cette fenêtre ses propres filtres. Pseudo-transparence veut dire que la conception du logiciel présuppose certains filtres. Le fait qu'avec DERIVE l'utilisateur obtienne des factorisations complètes dans Z ne pose pas problème aux mathématiciens qui sont les utilisateurs visés et qui conçoivent facilement qu'il y a d'autres factorisations, à la différence des élèves observés dans la seconde version. D'un point de vue didactique, le problème est que pour un enseignant non averti, les filtres sont ceux du mathématicien, d'où une étude trop superficielle du logiciel à disposition des élèves qui conduit à la non productivité de situations basées sur la simple visualisation comme dans la seconde version.

Un extrait d'interaction avec MistralAI (tableau 2) montre que la pseudo transparence persiste dans un système d'IA. Pour MistralAI il n'existe qu'une factorisation de $x^6 - 1$. Il répond à côté en mentionnant des factorisations sur d'autres anneaux. Après qu'une autre factorisation lui ait été proposée par l'utilisateur, le système admet l'existence de « factorisations valides mais pas complètes ». Sollicité à nouveau par la requête « toutes les factorisations sur Z de $x^6 - 1$ » MistralAI donne un « autre regroupement » $(x^2 - 1)(x^4 + x^2 + 1)$. Sollicité par plusieurs nouvelles requêtes insistant sur le nombre treize ($B_4 - 2$) de factorisations partielles possibles, et après une phase d'incohérence, MistralAI en donne 7 et persiste sur ce nombre.

Table 2 une interaction avec MistralAI. Les prompts sont soulignés.

<p><u>Je voudrais les factorisations de $x^6 - 1$ avec des coefficients entiers.</u></p> <p>Voici la factorisation de $x^6 - 1$ avec des coefficients entiers :</p> <p>Factorisation complète sur Z :</p> <p>$x^6 - 1 = (x - 1)(x + 1)(x^2 + x + 1)(x^2 - x + 1)$ Explication :</p> <p>Facteur linéaire : $(x - 1)$ et $(x + 1)$ proviennent des racines réelles $x=1$ et $x=-1$.</p> <p>Facteurs quadratiques : $(x^2 + x + 1)$ et $(x^2 - x + 1)$ proviennent des racines complexes j et j^2, où $j = e^{\frac{2i\pi}{3}}$.</p> <p><u>Est-ce que $(x^3 - 1)(x^3 + 1)$ n'est pas une factorisation de $x^6 - 1$?</u></p>
--

Tu as tout à fait raison, $(x^3 - 1)(x^3 + 1)$ est bien une factorisation de $x^6 - 1$ sur \mathbf{Z} !

Explication : On utilise l'identité remarquable $a^2 - b^2 = (a - b)(a + b)$...

Conclusion : $(x^3 - 1)(x^3 + 1)$ est une factorisation valide, mais elle n'est pas complète sur \mathbf{Z} .

La factorisation complète sur \mathbf{Z} est celle que j'ai donnée précédemment. Merci pour ta question, c'est une excellente remarque !

Toutes les factorisations sur \mathbf{Z} de $x^6 - 1$

...

Résumé des factorisations sur \mathbf{Z} :

Factorisation complète (irréductible) : $(x - 1)(x + 1)(x^2 + x + 1)(x^2 - x + 1)$

Factorisation intermédiaire : $(x^3 - 1)(x^3 + 1)$

Autre factorisation intermédiaire :

$$(x^2 - 1)(x^4 + x^2 + 1)$$

$$(x - 1)(x^5 + x^4 + x^3 + x^2 + x + 1)$$

....

Il y a 13 factorisations intermédiaires.

Pour $x^6 - 1$ il n'y a pas 13 factorisations intermédiaires distinctes sur \mathbf{Z} . Cependant, il existe plusieurs façons de regrouper les facteurs pour obtenir des factorisations intermédiaires. Voici les principales étapes et combinaisons possibles :

Factorisation complète (irréductible) :

$$(x - 1)(x + 1)(x^2 + x + 1)(x^2 - x + 1)$$

Factorisations intermédiaires possibles :

$$(x^3 - 1)(x^3 + 1)$$

$$(x^2 - 1)(x^4 + x^2 + 1)$$

$$(x - 1)(x^5 + x^4 + x^3 + x^2 + x + 1)$$

$$(x + 1)(x^5 - x^4 + x^3 - x^2 + x - 1)$$

$$(x^2 + x + 1)(x^4 - x^3 + x^2 - x + 1)$$

$$(x^2 - x + 1)(x^4 + x^3 + x^2 + x + 1)$$

$(x^3 - 1)(x^3 + 1)$ puis factoriser $(x^3 - 1)$ et $(x^3 + 1)$:

$$(x - 1)(x^2 + x + 1)$$

$$(x + 1)(x^2 - x + 1)$$

Autres combinaisons :

$$(x^2 - 1)(x^2 + x + 1)(x^2 - x + 1)$$

$$(x - 1)(x + 1)(x^4 + x^2 + 1)$$

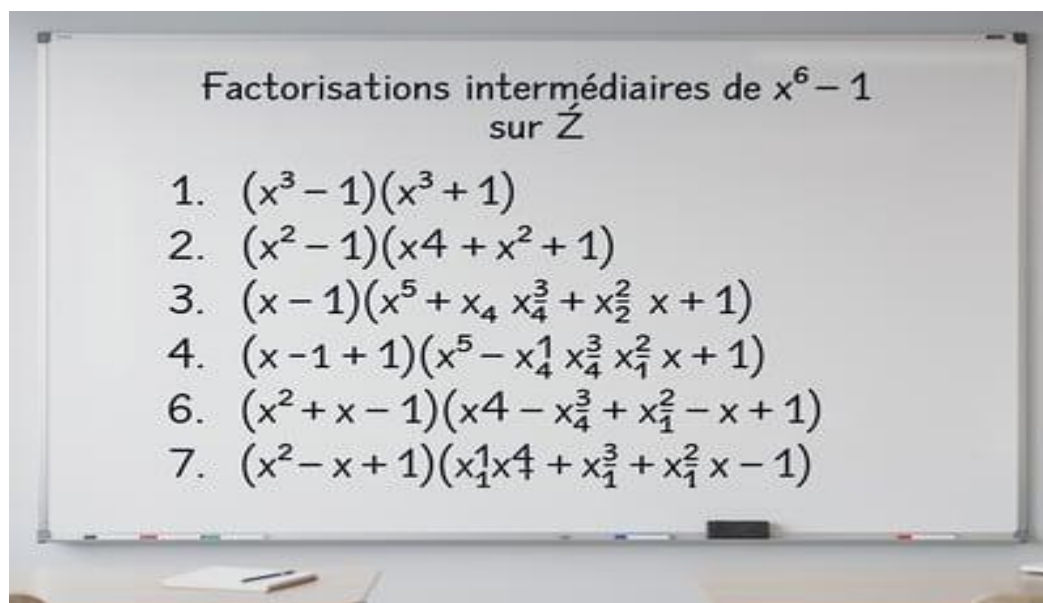
...

Ca fait 10, il en manque 3.

Ici MistralAI produit une liste avec 3 nouvelles factorisations, manifestement fausses car de degré 5, puis, après de nouveaux prompts devient incohérent et finalement se limite à une liste de 7 factorisations (6 en deux facteurs, 1 en trois facteurs).

Un dernier prompt : $(x^3 - 2x^2 + 2x - 1)(x + 1)(x^2 + x + 1)$

Voici une illustration des factorisations intermédiaires correctes de $x^6 - 1$ sur \mathbb{Z} :



Explication : La factorisation $(x^3 - 2x^2 + 2x - 1)(x + 1)(x^2 + x + 1)$ n'est pas correcte pour $x^6 - 1$, car le produit de ces facteurs ne donne pas $x^6 - 1$.

L'extrait d'interaction montre que des prompts adéquats permettent d'obtenir des factorisations partielles et donc de lever partiellement la pseudo-transparence. Cette procédure est différente de celle mise en œuvre sous l'impulsion des enseignants dans la troisième version avec DERIVE. Comment les enseignants vont-ils prendre conscience de la nécessité d'émettre des prompts adéquats ? Comment vont-ils préparer les élèves à les construire ? Ces questions interrogent la genèse instrumentale des ETM.

La fluidité du dialogue

Les requêtes en langage naturel permettent l'expression d'un besoin de façon moins contrainte qu'avec un logiciel de calcul formel et apportent une simplification par rapport aux équations dans les moteurs de recherche. Il s'agit du processus d'extraction de l'information, processus dans lequel les réseaux neuronaux excellent. Le champ des requêtes est élargi, par exemple la demande de factoriser $x^n - 1$ pour n entier est comprise par MistralAI, alors qu'elle n'a pas de sens pour un logiciel de calcul formel. Les réponses peuvent être qualifiées de polies, et même d'obséquieuses. Il s'agit de la capacité à engendrer du contenu caractéristique de ces IA. Un défaut repéré est le manque d'homogénéité, des réponses élémentaires (division par $x - 1$) coexistant avec d'autres à un niveau plus avancé (polynômes cyclotomiques). Quels sont les algorithmes sous-jacents ? Quels peuvent être leurs effets sur le travail mathématique ? Par exemple un élève curieux et doué ne sera pas sensible à l'hétérogénéité des réponses et prendra plaisir à mettre en défaut le système pour obtenir une excuse le valorisant, mais qu'en est-il pour un élève plus passif, quel jeu va-t-il jouer avec le système ? Il est possible de voir ces questions principalement comme intervenant dans la dimension sémiotique du travail.

Des erreurs (pudiquement dénommées hallucinations)

ChatGPT renvoie des factorisations manifestement fausses. Les erreurs de MistralAI sont plus subtiles. Le système déclare qu'il n'y a pas de factorisation radicale pour $x^5 - 1$, mais en

donne une pour $x^{17} - 1$. Dans la suite de l'interaction à propos des factorisations partielles de $x^6 - 1$ (tableau 2), MistralAI maintient qu'elles sont au nombre de 7, et déclare non valide la factorisation $(x^3 - 2x^2 + 2x - 1)(x + 1)(x^2 + x + 1)$ parce qu'elle n'appartient pas à la liste qu'il produit. Il finit par donner comme justification une image extraite d'un cours en ligne dont probablement il n'a retenu qu'une partie.

Il est clair que ChatGpt dans la version gratuite que j'ai utilisée n'a pas de calcul formel. MistralAI vérifie systématiquement les factorisations très probablement par calcul formel, mais, en dehors des tâches calculatoires, il est tributaire comme ChatGPT des pages web rencontrées au cours de son entraînement. Ces erreurs sont en quelque sorte les ratés du processus de raisonnement (dimension discursive). Elles conduisent à s'interroger sur le corpus ayant servi à l'entraînement et sur la fiabilité de cet entraînement. Les explications fournies pour une factorisation donnée laissent penser que pour les réponses les plus accessibles, le corpus est extrait de cours en ligne et de forums d'aide individualisée, et pour les réponses les plus techniques d'encyclopédies de mathématiques. Il est donc possible de comparer avec une compilation d'ouvrages, mais le défaut de cohérence n'existe-t-il pas nécessairement quand cette compilation est confiée à une machine ?

Vers un cadre théorique

Les dimensions discursive, sémiotique et bien entendu instrumentale interviennent dans le travail sur les factorisations. Le cadre des espaces de travail mathématiques est donc adapté. Cependant, Lagrange et Richard (2022) ont montré que la genèse instrumentale dans les ETM n'est pas de même nature que celle mise en avant dans l'approche instrumentale à laquelle se rattache la notion de *technique instrumentée*. Cette notion qui permet de rendre compte des différents processus à mettre en œuvre pour un travail mathématique avec des outils complexes comme le calcul formel et les LLM est donc à coordonner avec les différents éléments de la théorie des ETM.

Lagrange et Richard partent du constat d'une analyse souvent insuffisante de la genèse instrumentale dans la théorie des ETM. En général les deux pôles artefact et construction sont peu spécifiés et la genèse est assimilée à une version simplifiée de l'approche instrumentale. Cette version simplifiée se limite à postuler le développement de schèmes instrumentaux, entités psychologiques dont l'intégration dans la théorie des ETM aussi bien que leur observation dans les pratiques se révèlent peu évidents. Ainsi, l'assimilation des genèses de la théorie des ETM et de l'approche instrumentale appauvrit les deux théories. En revanche, clarifier les rapports entre les genèses de la théorie des ETM et celles de l'approche instrumentale peut être l'occasion de remettre l'approche instrumentale dans le champ des théories proprement didactiques loin des dérives psychologisantes qui sont apparues avec le temps.

Pour situer ces rapports, ce texte distingue d'une part les genèses « à la Rabardel » en se basant sur un article souvent cité, mais généralement de façon allusive (Vérillon et Rabardel, 1995) et d'autre part les genèses « à la Kuzniak » telles qu'introduites par l'auteur dans une présentation sur les espaces de travail géométriques à CERME 8 (Kuzniak, 2013).

Des genèses “à la Rabardel” aux techniques instrumentées

Les genèses “à la Rabardel” sont inspirées des genèses piagétienes qui s’opèrent par construction de schèmes, définis comme organisation des actions telles qu’elles se transfèrent ou se généralisent lors de la répétition de cette action. Concernant les savoirs, les genèses “à la Rabardel” se distinguent d’approches purement épistémiques pour rendre compte de situations où l’apprentissage s’inscrit dans un projet pratique et fonctionnel, par exemple l’usage d’un objet technique comme un tour ou un robot. Elles s’inspirent donc aussi de théorie de l’activité considérant que « le développement résulte de l’appropriation progressive de (...) réalisations socialement construites ». Dans les situations pratiques ces réalisations socialement construites sont des artefacts. Ceux-ci se définissent comme « les objets de la culture matérielle auxquels une personne a accès au cours de son développement ». Les artefacts interviennent dans l’activité instrumentée telle que résumée dans le schéma SAI (figure 1). L’artefact devient un *instrument* qui opère comme médiateur dans la relation du sujet avec l’objet de son activité. La médiation opère dans le sens Sujet-Objet quand l’artefact permet l’action sur l’environnement et dans le sens Objet-Sujet quand la médiation contribue à la conceptualisation par le sujet de son environnement.

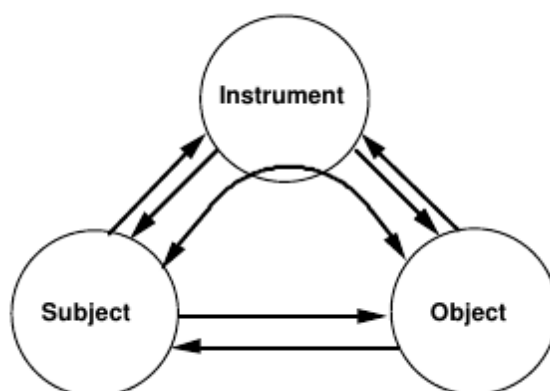


Figure 1 : Le diagramme SAI (Vérillon et Rabardel, 1995, page 85).

Je m’intéresse ici aux artefacts comme réalisations socialement construites et aux *techniques instrumentées* par lesquelles ils sont mis en œuvre. Cette *approche par les techniques* permet de dépasser une approche par les schèmes qui, si elle est utile du point de vue du sujet psychologique, atteint ses limites dans un contexte d’enseignement, ou plus précisément d’une institution ayant pour finalité un apprentissage donné. Cette approche par les techniques s’inscrit dans la théorie de l’activité mais il a été développé par Lagrange (2000, 2003) à partir de la notion de praxéologie de la théorie anthropologique du didactique (Chevallard, 1999). En effet, les artefacts, objets socialement situés, sont indissociables de techniques pour leur mise en œuvre elles aussi socialement situées. Dans l’enseignement, les institutions scolaires sont les lieux où les techniques trouvent leur légitimité en lien à la fois avec le niveau pratique des types de tâches qu’elles permettent d’accomplir et avec le niveau des théorisations où la cohérence et l’efficacité de ces techniques sont discutées. Les propriétés mathématiques, les concepts et un symbolisme spécifique apparaissent aussi à ce niveau des théorisations. L’apport d’une technique à un type de tâches est caractérisé selon Artigue (2002) par sa *valence pragmatique* alors que la contribution de la même technique à des théorisations

constitue sa *valence épistémique*. La valence pragmatique d'une technique est le plus souvent mise en avant pour justifier leur mise en œuvre. Si l'on considère le rôle du travail technique dans la construction des objets mathématiques (par exemple, on apprend ce qu'est une factorisation... en faisant des factorisations et en cherchant le meilleur moyen de les faire) la valence épistémique d'une technique est décisive pour l'apprentissage, mais elle est minorée dans la conscience des acteurs : une fois une théorisation installée, les moyens par lesquels cela s'est fait sont oubliés.

L'accent mis sur les techniques suppose une analyse fine des artefacts qu'elles mettent en œuvre. Dans quelles situations et pour quelles tâches un artefact donné peut-il être mis en œuvre ? Quelles sont précisément les caractéristiques de l'artefact qui vont permettre ces mises en œuvre ? Dans quelle mesure va-t-il entrer en concurrence avec des techniques habituelles légitimes dans l'enseignement ? Dans le cas par exemple de la calculatrice graphique, cet artefact va se trouver en concurrence avec les techniques papier/crayon pour l'étude des fonctions. Pour des tâches classiques comme l'étude d'une fonction polynôme de degré 3 ayant deux extrémaux et trois zéros dans \mathbf{Z} , la calculatrice graphique permet le repérage de ces valeurs par lecture sur l'écran et la production d'un tableau de variation de façon beaucoup moins laborieuse qu'une étude classique par dérivation et résolution d'équation. L'artefact est conçu pour résoudre de façon économique les cas les plus couramment rencontrés et se présente donc généralement avec une fenêtre initiale où les valeurs des deux coordonnées sont dans l'intervalle $[-10,10]$. La valence épistémique de l'étude d'une fonction dont les zéros et les extrema sont dans cette fenêtre (ce qui est généralement le cas dans les exercices scolaires) est ainsi limitée. Pour des fonctions qui échappent à ce standard, le repérage des valeurs critiques par lecture sur l'écran demande la mise en œuvre de techniques de cadrage. Ces techniques supposent des connaissances sur l'artefact (cadrage par zoom avant, arrière selon un des axes ou les deux, ou détermination explicite d'une fenêtre) et s'appuient aussi sur une forme d'étude de la fonction : il faut savoir où regarder. Leur valence épistémique est loin d'être négligeable. Ainsi l'analyse fine de la calculatrice permet de comprendre comment un artefact nouveau peut saper la valence pragmatique de techniques antérieures et rendre par conséquent obsolète les praxéologies auxquelles ces techniques participent. De nouvelles tâches et de nouvelles techniques sont à imaginer, en s'assurant que ces techniques ont à la fois la valence pragmatique qui justifie leur mise en œuvre et la valence épistémique qui fait qu'elles participent à la conceptualisation et ainsi, potentiellement, acquièrent une légitimité dans une institution d'enseignement donnée.

Un autre point important concerne spécifiquement les artefacts numériques. Pour comprendre cette spécificité, considérons les deux dispositifs dont la mise en œuvre par des sujets d'âge scolaire est étudiée par Vérillon et Rabardel, un tour et un robot, et comparons les genèses associées. Le tour sert à découper un matériau en une forme de révolution. Pour créer la forme souhaitée, le matériau est mis en rotation et l'opérateur déplace un outil de coupe le long de deux axes (respectivement parallèle et perpendiculaire à l'axe de rotation) à l'aide de vis visibles. Le robot sert à déplacer des objets dans un espace de travail en 3D. Extérieurement, le robot est composé d'un dispositif mécanique et d'un boîtier de commande. Le dispositif mécanique est similaire à un bras humain : une succession de trois segments mobiles relie une base fixe à une pince par quatre articulations motorisées. Le boîtier de commande est

composé de trois curseurs disposés comme les trois axes d'un système orthogonal XYZ. Entre le boîtier de commande et le dispositif se trouve un processeur programmé de telle sorte que chaque position des trois curseurs détermine une position de la pince dans une représentation cartésienne de l'espace de travail. Il en résulte que, contrairement au tour, il n'y a pas de relation directe visible entre l'action sur un curseur et le mouvement d'un élément isolé dans le dispositif. Le boîtier de commande est représenté par un triplet de coordonnées et le robot est représenté par la position de la pince dans l'espace. En interne, l'algorithme implémenté sur le processeur numérique utilise une autre représentation du robot, constituée d'une position angulaire pour chacune des quatre articulations, cohérente avec le fonctionnement mécanique du robot : pour un triplet de coordonnées donné, un algorithme calcule les valeurs des quatre positions angulaires afin d'assurer un positionnement adéquat de la pince. A la différence du tour, le robot se caractérise par des représentations de l'environnement et un traitement algorithmique. Selon cette caractérisation, le robot est un *artefact numérique* alors que le tour peut être qualifié de *dispositif analogique*.

Vérillon et Rabardel ont observé dans le cas du tour la genèse d'une relation géométrique entre le profil 2D créé par l'activation des vis et la forme 3D résultant du matériau en rotation. Avec le robot, la genèse est marquée par une prise de conscience très progressive des différentes représentations (internes et externes) de l'espace 3D et du traitement algorithmique entre le boîtier de commande et la partie mécanique, tel que programmé par le concepteur. Paradoxalement comme la partie mécanique est ouverte à l'observation, la représentation interne (position angulaire des quatre articulations) est plus visible que la représentation externe requise par l'interface utilisateur, en particulier pour les élèves dont la compréhension de l'espace 3D est à construire. C'est pourquoi les élèves associent initialement les actions sur les curseurs à des changements dans les positions angulaires des articulations. Ils sont surpris de ce que l'action sur un curseur entraîne le mouvement de plusieurs articulations. Ils associent néanmoins un petit mouvement d'un curseur avec un mouvement plus marqué d'une articulation. Ils tentent aussi de trouver une combinaison d'actions sur trois curseurs qui permettrait d'agir sur une seule articulation. Les élèves ne prennent conscience que progressivement du fait que la position de la pince est l'objet du contrôle par les curseurs, prenant implicitement conscience de la représentation externe du robot dans l'espace 3D et de l'algorithme à l'œuvre. Conformément à l'analyse de Vérillon et Rabardel, nous considérons le robot comme un instrument qui sert de médiateur dans la relation des élèves avec l'espace 3D, mais nous soulignons également le fait que l'artefact devient un instrument pour ces élèves grâce à une prise de conscience progressive des différentes représentations et de l'algorithme. La prise de conscience ne réside pas dans les détails, mais plutôt dans une compréhension globale de l'entrée et de la sortie, ainsi que de la nature de la transition entre celles-ci.

Considérons le cas d'artefacts plus généralement utilisés dans l'enseignement, la géométrie dynamique (GD) et le calcul formel (CF). Dans une thèse récente Lesnes-Cuisiniez (2022) reprend l'idée courante qu'un logiciel de GD *embarque* des connaissances géométriques mais il va plus loin en spécifiant comment ceci s'opère à travers le déplacement (*dragging*) des objets à la souris. Dans le cas des points, il distingue les points libres, les points libres sur objet et les points liés (« non attrapables »). Ces trois types de points ont des équivalents dans le langage géométrique (point quelconque dans le plan, etc.) mais en tant que représentation

dynamique leur emploi suppose une prise de conscience du graphe de dépendance qui lie les objets et de l'algorithme de mise à jour de la construction activé selon ce graphe lors d'un déplacement. À partir de là Lesnes-Cuisiniez distingue deux paradigmes propres à la GD : le paradigme des *constructions robustes* et le paradigme des *constructions molles*. Dans le premier paradigme une construction est effectuée selon un programme de création d'objets avec les fonctionnalités adéquates du logiciel (*outils*). Le graphe de dépendance est ainsi correctement constitué. S'il existe un ou des points libres, le déplacement de chacun de ces points va activer l'algorithme de mise à jour selon le graphe de dépendance, l'affichage à l'écran sera modifié, mais la construction conservera les propriétés géométriques inscrites dans le programme de construction. Dans une construction molle, l'utilisateur s'autorise en revanche à créer des points en les plaçant au jugé plutôt que par un outil de construction. Par exemple un point supposé au milieu d'un segment sera placé au jugé aligné et à égale distance des deux extrémités. Le point sera considéré comme un point non inclus dans le graphe de dépendance existant. Dans ce cas l'algorithme de mise à jour ne conservera pas la propriété de ce point et de ceux qui en dépendent.

Ces deux paradigmes correspondent en fait à deux techniques de construction qui se distinguent par leur valence pragmatique (plus élevée pour la construction molle) et leur valence épistémique (plus élevée pour la construction robuste). Concrètement dans les pratiques observées en classe, ces deux techniques sont rarement en évidence car la GD est plus souvent considérée comme un moyen commode d'obtenir une figure propre et des mesures (valence pragmatique) que comme un environnement dédié à la compréhension de la notion de construction et des propriétés géométriques associées (valence épistémique). La représentation de la construction par un graphe de dépendance et l'algorithme de mise à jour ne sont pas mobilisés. En revanche la recherche didactique sur la GD a mis en évidence le potentiel d'une dialectique entre les deux techniques comme le montre Leung (2014) à partir d'une synthèse des études dans ce domaine. La section suivante présentera un exemple.

Elbaz-Vincent (2005) caractérise le calcul formel comme un ensemble de représentations internes d'entités algébriques et d'algorithmes, en principe transparents pour l'utilisateur. Les logiciels de calcul formel offrent également une interface utilisateur avec une symbolique similaire à la symbolique mathématique habituelle. L'usage du calcul formel suppose cependant comme la GD une prise de conscience des algorithmes et représentations internes. Elbaz-Vincent donne l'exemple du traitement de l'expression $\cos(a)^2 + \sin(a)^2$ par cet artefact. Cette expression n'est pas *évaluée* à 1, même si a est déclaré réel. En revanche, elle est *simplifiée* à 1. La raison en est que l'*évaluation* et la *simplification* sont des algorithmes différents. Le premier effectue un traitement syntaxique standard afin d'obtenir une expression canonique, tandis que le second applique de manière heuristique un ensemble de transformations basées sur des règles et des propriétés mathématiques, dans le but de réduire la complexité de l'expression.

Quant au symbolisme, il n'est que partiellement similaire au symbolisme mathématique habituel, car les expressions canoniques doivent être compatibles avec les représentations internes et les algorithmes. Par exemple, les sommes, les limites et les intégrales et des calculs comme la factorisation et le développement doivent être saisis avec un symbolisme

fonctionnel pour le logiciel. De plus, le résultat correspond rarement directement à ce qu'attend l'élève, en particulier au niveau secondaire : c'est le cas, par exemple de $x^{\frac{3}{2}}$, lorsque l'utilisateur s'attend à $x\sqrt{x}$. Ceci rend les techniques d'utilisation du calcul formel particulièrement délicates à mettre en œuvre.

Genèses dans les ETM

Nous examinons maintenant les concepts fondamentaux de la théorie des ETM à la lumière de ce qui précède. Le premier concept est celui du travail mathématique. L'approche de Vérillon et Rabardel concerne principalement les domaines professionnels et techniques, comme le montrent les études empiriques (tour, robot) citées dans l'article, et repose sur l'idée d'activité plutôt que de travail. Néanmoins, modéliser un objet sur un tour, activer un robot pour une tâche donnée, ne sont pas des activités libres menées dans le but de développer des connaissances géométriques ou informatiques abstraites, mais plutôt des activités visant à la transformation d'entités matérielles et se développant dans des contextes sociaux. De son côté la théorie des ETM, en mettant l'accent sur le *travail* et en considérant les mathématiques du point de vue du travail qu'effectuent mathématiciens et apprentis mathématiciens rompt avec l'idée des mathématiques comme domaine abstrait déconnecté de pratiques sociales : comme l'opérateur de tour ou de robot, le mathématicien travaille avec des instruments qui ont une valeur psychologique et institutionnelle.

La genèse instrumentale, sur laquelle nous nous intéressons particulièrement ici, relie les *artefacts*, un pôle du plan épistémologique, et la *construction*, le pôle correspondant dans le plan cognitif. Sur le diagramme classique des ETM les genèses semblent à première vue avoir un caractère objectif, puisque les composantes du travail mathématique interviennent dans des processus observables : la genèse instrumentale avec une équerre, par exemple, serait simplement l'utilisation de l'équerre pour tracer des angles droits dans une construction donnée. Néanmoins, Kuzniak insiste sur le caractère *développemental* des genèses, suggérant que des transformations ont lieu au sein du ou des sujets ou des institutions qui utilisent les artefacts. Kuzniak (2022) distingue deux points de vue, ascendant et descendant, sur les genèses. En particulierisant à la genèse instrumentale, le point de vue ascendant décrit les actions par lesquelles l'utilisateur s'approprie les différentes techniques d'utilisation de l'artefact. Le point de vue descendant part de la réalisation souhaitée par l'utilisateur pour aboutir au choix adéquat de l'artefact avec lequel il va effectuer les actions requises. Ainsi la genèse instrumentale modifie la manière dont l'utilisateur agit et pense, et, en retour, transforme également les connaissances mathématiques. La genèse instrumentale dans les ETM est donc compatible avec l'idée de médiation instrumentale de Vérillon et Rabardel (diagramme SAI fig. 1), mais à la différence de l'approche instrumentale, elle ne théorise pas les dimensions liées au sujet (agir et penser) et aux institutions (techniques instrumentées, impact sur les praxéologies).

Une autre différence est le champ d'application de la genèse instrumentale : Vérillon et Rabardel, se référant aux instruments psychologiques de Vygotsky, expliquent dans leur conclusion que leurs remarques ne valent pas seulement pour les activités instrumentalisées par des artefacts matériels : « la dimension instrumentale concerne également la maîtrise du

langage, de l'écriture et des chiffres, qui peuvent être considérés comme des technologies destinées à communiquer, à représenter et à conserver des informations, et à calculer ».

Ainsi, Vérillon et Rabardel après avoir regretté que les processus psychologiques liés aux artefacts techniques soient négligés par la psychologie élargissent dans leur conclusion le champ d'application à la genèse de tout type d'artefact. L'idée de genèse instrumentale perd ainsi une partie de son potentiel en étant absorbée dans une catégorie trop large de processus cognitifs comme on peut le constater aujourd'hui dans beaucoup de recherches qui se revendiquent de l'approche instrumentale. En revanche, le plan épistémologique des ETM distingue trois pôles, chacun pouvant être considéré comme un artefact dans une genèse spécifique. La théorie des ETM apporte donc une clarification, nous aidant à considérer la spécificité de la genèse instrumentale.

Pour aller plus loin, il convient de caractériser les artefacts du pôle instrumental. Dans le plan épistémologique des ETM, les objets de chaque pôle peuvent avoir une base matérielle aussi bien qu'immatérielle. Cela est évident pour les signes dans le pôle sémiotique. Quant au corpus de référence dans le pôle discursif on les voit souvent comme un ensemble d' « idées » dépourvues de base matérielle mais Richard et al (2018) précisent que cet ensemble est souvent « matérialisé » de diverses manières (recueil de propriétés dans un cours papier/crayon, règle de déduction dans un système tutoriel...). La forme matérielle des instruments traditionnels (règle compas...) ou des logiciels contemporains ne doit pas cacher que leur réalisation physique découle de propriétés ancrées dans des théories et des institutions et, pour les logiciels, implémentées sous forme de représentations et d'algorithmes.

Finalement la bonne manière de caractériser les objets du plan épistémologique est de considérer leur rôle dans la genèse correspondante. Donc les artefacts instrumentaux sont ceux qui interviennent dans une *construction*. Il faut alors préciser ce qu'on entend par construction et pour cela il est utile de revenir sur la définition proposée par Duval (1995) : « (une) construction (...) peut fonctionner comme un modèle en ce sens que les actions sur le représentant et les résultats observés sont liés aux objets mathématiques qui sont représentés. » La construction est donc l'élaboration de modèles de configurations mathématiques c'est-à-dire d'un ensemble d'objets interdépendants, plutôt qu'un ou plusieurs objets isolés. Cela correspond bien aux figures géométriques traditionnelles qui modélisent les propriétés d'un ensemble d'objets géométriques : le mathématicien géomètre va tracer une ou des figures lui permettant d'explorer des propriétés. Les logiciels ajoutent une dimension dynamique, donnant un caractère de simulation au modèle. En algèbre, une calculatrice numérique aide à faire fonctionner une formule comme modèle d'une fonction en fournissant des valeurs calculées pour des entrées données. En interprétant l'étude de Vérillon et Rabardel sur les robots dans la théorie des ETM, nous devons comprendre que la construction est ce que font les élèves lorsqu'ils font fonctionner le robot à travers les microgenèses. Tout comme une figure en géométrie n'est qu'une des réalisations possibles d'une configuration, une action sur les curseurs n'est qu'une des actions possibles. Cependant, on sait que dessiner une figure aide à penser toutes les réalisations possibles, c'est-à-dire qu'elle est considérée comme un modèle de la configuration. De même, les actions des élèves sur les curseurs sont un moyen de tester un modèle implicite du robot et de faire évoluer ce modèle. Cela met en lumière une

caractéristique de la genèse instrumentale : il ne s'agit pas d'une application passive de techniques standard conformes à un mode d'emploi, mais plutôt de l'élaboration active d'une technique adaptée en vue d'obtenir des résultats pertinents. Cette élaboration participe au travail sur les techniques souligné par Chevallard (1999).

Soulignons à ce propos que la théorie des ETM et l'approche instrumentale divergent quant au statut des techniques dans la genèse. Selon Kuzniak et al (2016) « les artefacts sont intimement liés aux règles et techniques de construction ou de calcul (division euclidienne, constructions classiques à la règle et au compas, etc.). Ces techniques reposent sur des algorithmes dont la validation et le statut théorique ne posent plus de problème. » L'affirmation que je souligne n'est pas compatible avec l'approche instrumentale puisque celle-ci postule que la validation et le statut des techniques sont précisément ce qui est en jeu dans la genèse. La routinisation sous forme d'un algorithme est postérieure à ce travail sur les techniques et tend à l'oblitérer. Cette analyse a été développée dans la recherche sur les usages du calcul formel dans les années 1990 face un courant dominant notamment nord-américain qui voyait le calcul formel comme un moyen de minorer le travail technique, supposé routinier et peu producteur de conceptualisations, de façon à passer directement des tâches aux théorisations (Lagrange 2000, Artigue 2002).

Un exemple

Voici un exemple de l'apport d'une analyse des techniques instrumentées, à une situation étudiée dans le cadre des ETM par Kuzniak (2013). Les élèves de seconde (10ème grade) sont confrontés à la tâche de construire en papier crayon un triangle ayant les mêmes angles qu'un triangle donné (c'est-à-dire un triangle semblable) puis d'émettre des conjectures sur les longueurs des côtés des deux triangles. L'enseignant part du principe que la diversité des triangles obtenus leur permettra d'inférer la proportionnalité entre les longueurs des côtés du triangle initial et celles du triangle construit. Il vise à institutionnaliser la proportionnalité sans démonstration, mais il souhaite néanmoins que les élèves soient conscients qu'il s'agit d'une conjecture non prouvée, l'imprécision des mesures permettant cette prise de conscience. Kuzniak illustre ces attentes de l'enseignant dans un premier diagramme qui ne mentionne aucune genèse instrumentale. Les artefacts et la construction sont liés à la genèse sémiotique : ils « servent » simplement la visualisation, et leur lien n'est pas discuté. La construction est également censée « déclencher » une genèse discursive.

L'absence de la genèse instrumentale dans l'anticipation du professeur s'analyse comme une sous-estimation de la dimension technique de la tâche demandée aux élèves. Il s'agit d'une tâche de report d'angle. La technique la plus efficace serait celle du papier calque avec des variantes utilisant un rapporteur ou du papier plié. Ces techniques appartiennent au répertoire des constructions du cycle 4 et il serait étonnant qu'elles soient mobilisées par des élèves de lycée. À supposer que les élèves disposent d'un rapporteur, son maniement demande une habitude certainement perdue au lycée. La technique canonique au lycée serait le report au compas, mais on ne peut pas s'attendre à ce qu'elle soit élaborée spontanément.

Après que les élèves n'ont pas réussi la tâche de construction ou n'ont pas réussi à s'engager dans un processus de conjecture, l'enseignant présente une figure de géométrie dynamique

(GD) sur un ordinateur, fait calculer par le logiciel les rapports entre les mesures des côtés pour diverses configurations, puis demande aux élèves d'observer l'invariance des rapports. La figure de géométrie dynamique permet aux élèves de passer « directement » de la visualisation à une « preuve » empirique basée sur l'observation de l'invariance pour plusieurs configurations. La précision des mesures effectuées à l'aide du logiciel et la pluralité des configurations font croire aux élèves qu'ils ont réalisé une « véritable démonstration », ce qui n'était pas l'objectif de l'enseignant. Là encore, il n'y a aucune genèse instrumentale.

Ainsi le professeur clôt la phase de construction et de conjecture par les élèves malgré leur échec dans la tâche. Ceci s'explique si l'on considère la valence épistémique des différentes techniques de report d'angle. Cette valence épistémique est relative à la notion d'angle qui n'est pas l'objectif de la séance et un professeur ne peut pas se permettre de « perdre du temps » dans un travail sur des techniques dont la valeur épistémique ne correspond pas à son objectif.

Le choix de la GD par le professeur dans la seconde phase découle de la valence pragmatique des techniques d'utilisation de la GD : construction robuste (au sens précisé ci-dessus) qui permet de faire varier les configurations et visualisation dynamique des mesures et de leur rapport. Il est seul à mettre en œuvre ces techniques car ici aussi la valence épistémique de ces techniques ne l'intéresse pas. Il vise l'efficacité, au prix d'une confusion chez les élèves entre conjecture et preuve. Le travail sur les techniques de GD qui aurait permis de faire la différence entre ce qui est exact (la construction) et ce qui est approché (les mesures) a été négligé.

Une reconstruction

Une attention au travail sur les techniques permet d'envisager une alternative à la mise en œuvre par le professeur. Il s'agit d'imaginer l'élaboration d'une technique de construction d'un triangle semblable qui mette en évidence l'équivalence de la propriété d'égalité des angles et l'égalité des rapports. La reconstruction utilise la dialectique construction molle-construction robuste présentée plus haut. Concrètement, un triangle ABC et un triangle DEF sont créés en géométrie dynamique. Les points A, B et C sont libres dans le plan. D et E sont fixes. Il y a ensuite trois phases (Figure 2).

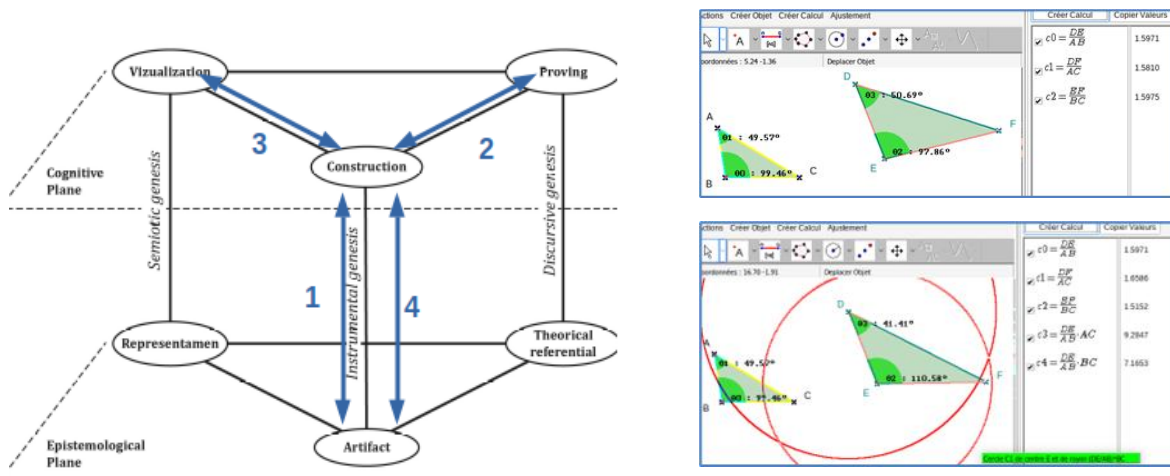


Figure 2 les techniques de construction dans l'espace de travail de la situation reconstruite

1. **Construction molle à l'aide d'un point libre.** Appropriation de la figure et ajustement à la souris du triangle DEF de façon qu'il ait les mêmes angles que ABC. Cette phase de construction molle est facilitée par l'affichage de valeurs approchées des angles dans le logiciel. Point fixes et points libres sont distingués. Les élèves prennent conscience de ce que l'égalité des angles ne résiste pas au déplacement d'un des sommets du triangle ABC.
2. **Prise de conscience d'une construction qui ne résiste pas au dragging.**
3. **Visualisation de la proportionnalité des longueurs pour s'engager dans une construction robuste.** Le triangle DEF restant en construction molle, on peut demander de créer des points E' et F' respectivement sur (DE) et (DF) pour que DE'F' et ABC soient perceptivement isométriques. Ainsi, la proportionnalité peut émerger à partir de la propriété de Thales.
4. **Construction robuste à l'aide de cercles et d'intersections.** Construction robuste à l'aide de cercles et d'intersections. Les distances DF et EF sont calculées par le logiciel à partir de l'égalité des rapports. Le point F peut alors être construit comme une des deux intersections de cercles respectivement de centre D et E et de rayon les distances calculées.

Les techniques de construction molle s'inscrivent dans une exploration perceptive, ce qui montre bien la proportionnalité comme une conjecture. Sans constituer une démonstration, la construction robuste met en évidence la relation de proportionnalité comme un outil utile pour la construction plutôt qu'un résultat à simplement ajouter au répertoire de référence. La reconstruction est résumée dans la fig. 2 qui montre les phases de construction dans la genèse instrumentale.

Le tableau 3 résume l'analyse des techniques intervenant dans la situation initiale et dans la reconstruction proposée. Une technique par report d'angle est supposée par le professeur être une routine pour les élèves. Elle n'est pas commode (Valence Pragmatique faible), elle n'apporte pas d'éléments pour la conceptualisation visée dans la séance (Valence Epistémique faible), mais elle est reconnue dans l'institution (Viabilité institutionnelle). Avec la GD, les techniques de construction molle et robuste s'opposent quant à leurs valence pragmatique et épistémique et à leur viabilité institutionnelle. La construction molle a une valence heuristique, mais n'est pas destinée à être une procédure canonique. La construction robuste est plus difficile à mettre en œuvre, mais elle fait intervenir des connaissances mathématiques, particulièrement celles qui sont en jeu dans la situation. De plus, la construction robuste est un enchaînement de primitives géométriques qui peut se faire reconnaître par l'institution comme l'équivalent d'une construction à la règle et au compas.

Table 3 Valences et viabilité des techniques de construction géométrique

Technique	Valence Pragmatique	Valence Epistémique	Viabilité institutionnelle
Report d'Angle	Faible	Faible	Oui (routine PC)
Construction Molle	Forte	Faible	Non
Construction Robuste	Moins Forte	Plus Forte	Oui (à construire)

A partir des situations sur les factorisations avec le calcul formel, le tableau 5 montre une analyse des techniques analogue à celle ci-dessus (tableau 4) dans le cas de la GD. Une figure

analogue à la figure ci-dessus (fig. 3) résume le parcours dans les différentes genèses et la place centrale de la genèse instrumentale des techniques dans la troisième version.

Table 4 Valences et viabilité des techniques de factorisation

Technique	Valence Pragmatique	Valence Epistémique	Viabilité institutionnelle
Division	Faible	Faible	Oui (routine PC)
CF presse-bouton	Forte	Faible	Non
CF presse-bouton+ regroupement des facteurs	Moins Forte	Plus Forte	Oui (à construire)

Synthèse : la théorie des ETM et les techniques instrumentées

Il serait vain de chercher à fusionner deux approches théoriques issues, comme montré ci-dessus, de perspectives théoriques différentes. Il est préférable de souligner les apports réciproques et les divergences. L'analyse et l'exemple qui précèdent montrent comment l'approche instrumentale peut étoffer la genèse instrumentale des ETM. L'accent mis sur les techniques instrumentées et leurs valences pragmatique et épistémique « colle bien » avec l'idée de travail, centrale dans les ETM. Cela tranche avec une analyse limitée à des schèmes instrumentaux peu opératoire à elle seule pour l'analyse des situations d'enseignement. En retour, la théorie des ETM apporte à l'approche instrumentale en distinguant trois genèses permettant ainsi d'identifier ce qui relève réellement d'une dimension instrumentale et l'articulation avec les dimensions symboliques et discursives, les trois dimensions étant trop souvent confondues dans les analyses se réclamant de l'approche instrumentale. Nous avons souligné aussi que, à la différence de la théorie des ETM, l'approche instrumentale met l'accent sur le travail sur les techniques plutôt que sur la mobilisation de techniques routinisées.

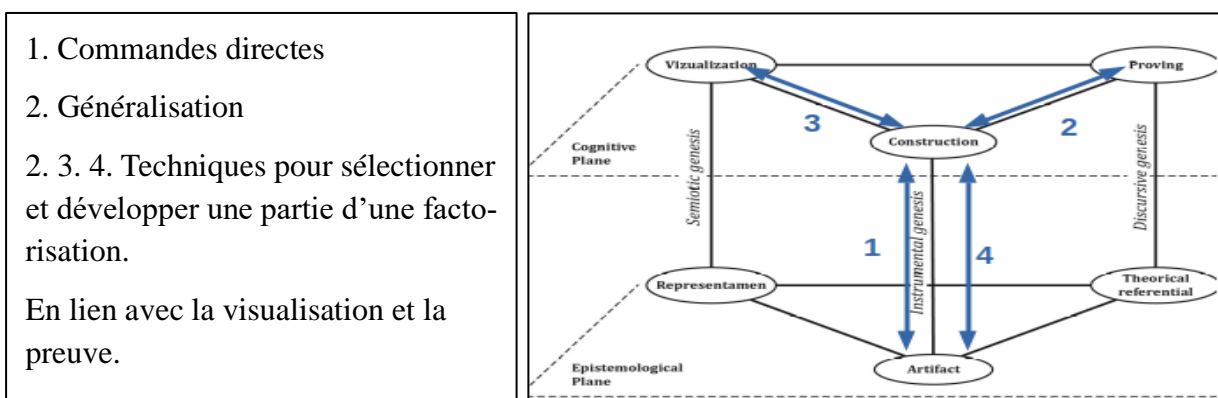


Figure 3 : Les techniques de factorisation dans l'espace de travail de la troisième version des factorisations.

Une analyse des artefacts

Dégager des techniques d'usage d'artefacts numériques, étudier leurs valences et imaginer des situations didactiques où elles pourraient émerger suppose une analyse de ces artefacts qui ne se limite pas à des potentialités générales. Il s'agit, sur un type de tâche donné de comprendre de façon suffisamment précise quels algorithmes et représentations sont à l'œuvre. Je

commence par la factorisation des polynômes pour montrer comment une telle étude peut être menée. Puis j'essaie de dégager quelques pistes pour une étude similaire dans le cas des LLM.

Données sur la factorisation des polynômes à une variable par le calcul formel

Comment les logiciels de calcul formel factorisent-ils les polynômes à une variable. Si un polynôme Q de degré d divise un polynôme P , alors tous les coefficients de Q sont majorés en valeur absolue par $2^d \|P\|$, la norme $\|P\|$ étant la somme des valeurs absolues des coefficients de P (inégalité de Mignotte-Landau). Une procédure simple consisterait donc à explorer de façon exhaustive les polynômes de degré inférieur à celui de P dont les coefficients respectent l'inégalité. Il est facile de se rendre compte de ce qu'une telle procédure appliquée telle quelle serait irréaliste en dehors de cas très simples. C'est pourquoi des algorithmes ont été inventés de façon à réduire l'exploration, notamment en passant par des factorisations dans $\mathbf{Z}/n\mathbf{Z}$ pour n bien choisi et en utilisant d'autres majorations.

Le cas des polynômes de la forme $x^n - 1$ est particulier car il existe un algorithme déterministe qui garantit une factorisation complète dans \mathbf{Z} . En effet, dans \mathbf{C} , $x^n - 1$ a n racines, les racines $n^{\text{ièmes}}$ de l'unité. La factorisation dans \mathbf{C} est donc triviale. La factorisation dans \mathbf{R} sous forme de produits de trinômes s'opère en combinant les binômes conjugués. Elle n'est généralement pas donnée par les logiciels de calcul formel, sauf lorsque $\cos\left(\frac{2p}{n}\right)$ s'exprime avec des racines carrées (factorisation radicale). C'est le cas lorsque n est un nombre de Fermat $n = 2^{2^k} + 1$. Le tableau 5 donne un exemple où un logiciel (Xcas) retourne la factorisation radicale pour le premier nombre de Fermat (5), mais pas pour le second (17). Il peut y avoir plusieurs explications. (1) Le logiciel a une méthode pour $x^5 - 1$ passant par la résolution d'équations du second degré, mais pas pour $x^{17} - 1$. (2) les expressions radicales des coefficients des trinômes factorisant $x^{17} - 1$ sont très volumineuses et c'est un choix du concepteur de ne pas retourner ce type d'expression. En général les logiciels de calcul formel retournent par défaut une factorisation dans \mathbf{Z} et il faut préciser une commande spéciale pour les factorisations dans \mathbf{C} ou \mathbf{R} . Le concepteur de Xcas opère donc un choix spécifique qui conduit à une factorisation dans $\mathbf{Q}[\sqrt{5}]$ pour $x^5 - 1$ et dans \mathbf{Z} pour les autres polynômes.

Table 5 Les factorisations avec un logiciel de calcul formel.

Tapez les instructions XCAS ci-dessous ou aidez-vous du clavier en haut à droite	
1	factor ($x^5 - 1$) $(x - 1)\left(x^2 + \frac{-(\sqrt{5}) + 1}{2} \cdot x + 1\right)\left(x^2 + \frac{\sqrt{5} + 1}{2} \cdot x + 1\right)$
2	factor ($x^9 - 1$) $(x - 1)(x^2 + x + 1)(x^6 + x^3 + 1)$
3	factor ($x^{12} - 1$) $(x - 1) \cdot (x + 1)(x^2 + 1)(x^2 - x + 1)(x^2 + x + 1)(x^4 - x^2 + 1)$
4	factor ($x^{17} - 1$) $(x - 1)(x^{16} + x^{15} + x^{14} + x^{13} + x^{12} + x^{11} + x^{10} + x^9 + x^8 + x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x + 1)$

Pour comprendre comment les logiciels de calcul formel opèrent la factorisation dans \mathbf{Z} des polynômes de la forme $x^n - 1$, partons de la factorisation dans \mathbf{C} , c'est à dire comme produit des binômes $(x - \alpha^k)$, ($k = 0..n - 1$), les α^k étant les racines n -ièmes de l'unité. Pour chaque diviseur d de n , on considère les $E_d = \{k \mid 0 < k \leq n \text{ et } \text{pgcd}(k, n) = d\}$ et les polynômes $\phi_d(x)$ produits des $(x - \alpha^k)$ tels que k appartient à E_d . Les E_d réalisent une partition de l'intervalle de $\mathbf{N} [1, n]$, et donc aussi les E_d . Il en résulte que $x^n - 1$ est le produit des polynômes $\phi_d(x)$. On démontre que $\phi_d(x)$ est le produit des binômes $(x - e^{\frac{2i\pi k}{d}})$ pour $0 < k \leq d$ et $\text{pgcd}(k, d) = 1$. Il ne dépend donc que de d est appelé $d^{\text{ième}}$ polynôme cyclotomique. Il est démontré que les polynômes cyclotomiques sont irréductibles dans \mathbf{Q} et à coefficients entiers.

Dans la deuxième ligne de l'exemple du tableau 2, $n=9$ a trois diviseurs, 1, 3 et 9 :

$$\phi_1(x) = (x - 1) \text{ correspond au diviseur } d = 1,$$

$$\phi_3(x) = x^2 + x + 1 = (x - e^{2i\frac{\pi}{3}})(x - e^{4i\frac{\pi}{3}}) \text{ correspond au diviseur } d = 3,$$

$$\phi_9(x) = x^6 + x + 1 = (x - e^{2i\frac{\pi}{9}})(x - e^{4i\frac{\pi}{9}})(x - e^{8i\frac{\pi}{9}})(x - e^{10i\frac{\pi}{9}})(x - e^{14i\frac{\pi}{9}})(x - e^{16i\frac{\pi}{9}}) \text{ correspond au diviseur } d = 9.$$

Les logiciels de calcul formel utilisent $\phi_1(x) = (x - 1)$ et la formule $\Phi_n(x) = \frac{x^n - 1}{\prod_{d|n, d < n} \phi_d(x)}$ pour calculer par récurrence les polynômes cyclotomiques et de là une factorisation de $x^n - 1$ sous forme de facteurs irréductibles dans \mathbf{Q} . Diverses propriétés des polynômes cyclotomiques (par exemple $\Phi_{2n}(x)$ est égal à $\Phi_n(-x)$ si n est impair, et à $\Phi_n(x^2)$ si n est pair) sont utilisées pour une plus grande efficacité.

Pour une situation de classe basée sur les factorisations de $x^n - 1$ par le calcul formel, les conséquences sont les suivantes :

- Les élèves obtiendront par la simple émission d'une commande une factorisation pour une valeur de n choisie (technique presse-bouton).
- Généralement, les élèves obtiendront une factorisation dans \mathbf{Z} . Il est possible qu'ils obtiennent une factorisation radicale dans certains cas, soit par décision du concepteur comme dans le tableau 2, soit parce qu'ils auront activé cette factorisation optionnelle.
- Les factorisations seront complètes, au sens de produits de facteurs irréductibles. En particulier la factorisation générique $(x - 1)(x^{n-1} + \dots + x + 1)$ ne sera obtenue que pour n premier, et la factorisation $x^{2n} - 1 = (x^n - 1)(x^n + 1)$ ne sera obtenue que pour $n=1$.

Données sur le fonctionnement des LLM rendant compte de l'interaction pour la factorisation des polynômes

Je rappelle trois caractéristiques relevées empiriquement dans le domaine des factorisations, et les questions qu'elles posent : (1) la pseudo-transparence et les techniques pour y parer (2) les algorithmes cachés pilotant les réponses du système et leur influence sur la façon dont les

élèves peuvent interagir avec le logiciel (3) le défaut de cohérence dans les sources servant à l'entraînement du système, et la nécessité pour les élèves d'en prendre conscience et de l'intégrer dans leurs techniques d'interaction.

Est-il possible d'avoir une analyse qui dépasserait ce type de constat empirique au cas par cas ?

Je ne cherche pas à me situer dans un débat sur l'intelligence des machines et des humains, mais plutôt à baser mon analyse sur les principes de développement des IA. Les capacités de l'IA générative sont le plus souvent présentées comme découlant d'une approche statistique : la recherche du token (unité élémentaire d'information, par exemple un mot dans la génération de texte) le plus probable complétant une suite de tokens donnée grâce une matrice de proximité de très grande dimension. Comme beaucoup d'autres auteurs, Aghion et Bouverot (2025) décrivent ainsi cette approche : « à chaque possibilité de token successeur est associée une probabilité. La réponse est générée, token par token, en fonction de la probabilité de chaque token. Ce fonctionnement présente une caractéristique essentielle : les modèles d'IA générative n'appréhendent pas la signification des mots, des images ou du son. » Pour moi, cette caractérisation, qui réduit les LLM à des algorithmes de complétion de phrase (comme ceux des téléphones mobiles) ne dit rien sur l'IA générative. Qu'est-ce qu'appréhender une signification, si ce n'est créer une structure au plus profond d'un réseau neuronal ? L'IA générative ne se contente pas d'aligner des token sur un fragment de phrase. Une série de représentations et de traitements complexes permet la création de structures dont seuls les effets sur la génération de contenu peuvent s'observer. Différentes *fenêtres de contexte* permettent la prise en compte à la fois de tokens proches et d'autres plus éloignés. Si l'on désigne par *apprentissage* un tel processus et *généralisation* la création de contenus en réponse à un prompt non connu à l'avance, alors ces systèmes apprennent dans des domaines ou sur des tâches données, sans programmation explicite, et sont capables de généraliser. Les traitements interviennent a priori pour créer les structures et a posteriori pour corriger les comportements dans une finalité *d'alignement* avec les objectifs de leurs concepteurs. Dans l'ensemble ces traitements sont opaques.

La question n'est pas alors de savoir si ces systèmes appréhendent ou non des significations, mais plutôt comment et dans quelle mesure ils réalisent ce pourquoi ils ont été conçus. Il serait possible de répondre de façon sarcastique. Oui, ils s'expriment de façon correcte et fluide, avec une tendance à la flagornerie, mais ce n'est pas une raison pour prendre leur discours comme argent comptant. Certes ils donnent l'apparence d'un apprentissage, mais plutôt comme un élève très moyen : connaissances peu assurées, références hétérogènes. Certes ils généralisent, mais pas toujours à propos.

Parmi les processus mis en œuvre par les IA génératives dans le cas des factorisations, les processus de raisonnement et d'explication sont les points faibles, comparés à l'extraction d'information et à la production de contenus. De plus, ils ont un caractère plutôt chaotique, de légères variations dans la requête pouvant provoquer de grandes différences dans les réponses. Les différences s'observent aussi d'un LLM à l'autre. Est-ce à dire qu'un travail avec les élèves ne présenterait pas d'intérêt ? Mon expérience de recherche de toutes les factorisations de $x^6 - 1$ m'incite à penser le contraire. Les réponses et arguments avancés par l'IA et la

nécessité de réfutations m'ont amené à réactiver mes connaissances un peu lointaines sur les polynômes, en lien avec l'acquisition de connaissances sur le nombre de partitions d'un ensemble à 4 éléments. Bref, mon intuition est que l'on peut apprendre en interagissant avec un LLM sous réserve de construire des techniques adaptées reposant elles-mêmes sur une certaine compréhension du fonctionnement du système de disposer de connaissances permettant d'appréhender des connaissances nouvelles. Cette intuition est confortée par des retours d'enseignants experts comme Durand (2025).

Pour appuyer cette intuition, voici quelques éléments inspirés de commentaires de Victor Storchan (<https://legrandcontinent.eu/fr/2026/01/28/lia-est-un-risque-existential-lalerte-de-dario-amodei-texte-integral-comment/>).

Les limites des LLM

Storchan suggère que des « limitations » perçues des systèmes d'IA ne sont en fait pas des barrières intrinsèques, mais des objectifs temporaires liés à ce que à ce qu'il est choisi de mesurer à un moment donné. Il prend l'exemple de *benchmarks* en mathématiques :

« GSM8K⁴ mesurait surtout la capacité à enchaîner correctement des opérations arithmétiques de niveau scolaire ; avec l'augmentation de l'échelle et l'usage du raisonnement pas à pas (chaînes de pensées)⁵, il a été rapidement saturé. AIME⁶ a déplacé l'objectif vers des problèmes de concours, plus courts mais exigeant des intuitions algébriques et combinatoires, où la progression est plus lente et irrégulière. D'autres *benchmarks* plus récents comme FrontierMath⁷ évaluent la capacité à maintenir un raisonnement mathématique cohérent sur des problèmes longs et peu standardisés, et restent aujourd'hui non saturés. »

Selon Storchan, dès qu'une faiblesse est clairement formalisée par un *benchmark* elle attire des efforts de collecte de données, d'entraînement et d'optimisation, et l'écart de performance se réduit rapidement.

Je vois trois conséquences : (1) il est périlleux de se positionner de façon générale sur l'usage d'IA génératives en fonction de limites constatées à un moment donné et qui, possiblement, vont être dépassées. (2) En revanche, pour une situation d'enseignement donnée, il sera utile d'explorer le *benchmark* ayant servi à l'évaluation du ou des LLM à disposition des élèves, notamment les limites de ces benchmarks. Par exemple les modèles entraînés sur GSM8K sont sensibles à certaines variations de l'énoncé notamment l'introduction de données non

⁴ The GSM8K benchmark comprises 1,319 grade school math word problems, each crafted by expert human problem writers. These problems involve elementary arithmetic operations (+ - × ÷) and require between 2 to 8 steps to solve.

⁵ Chain-of-Thought (CoT) prompting is an approach where the model is prompted to articulate its reasoning process to arrive at an answer.

⁶ All 30 problems from the 2025 American Invitational Mathematics Examination, testing olympiad-level mathematical reasoning with integer answers from 000-999.

⁷ *FrontierMath* is a benchmark of hundreds of original, exceptionally challenging mathematics problems crafted and vetted by expert mathematicians.

pertinentes⁸. Plus généralement, il est possible que les benchmarks reflètent les représentations communément admises à partir des pratiques scolaires, un peu comme les calculatrices graphiques offrent une fenêtre initiale x et y allant de -10 à 10 reflétant le fait que la grande majorité des fonctions étudiées dans le cadre scolaire ont leurs propriétés dans cette fenêtre. Cela peut conduire à construire des situations didactiques où les élèves auraient à dépasser ces limites et qui pourraient fonctionner tant que de nouvelles versions entraînées sur des benchmarks améliorés ne seront pas disponibles. (3) Il reste à déterminer s'il y a des limites inhérentes, communes avec le calcul formel et liées à la pseudo transparence comme la supposée unicité de la factorisation.

Les modèles de raisonnement

Pour Storchan, « les modèles de raisonnement ne tirent pas leur efficacité uniquement d'un calcul plus long à l'inférence mais également de leur capacité à simuler implicitement des interactions complexes de type multi-agents dans laquelle différentes perspectives internes, associées à des traits de personnalité et à des expertises distinctes, se confrontent, débattent et se complètent pour produire un raisonnement plus robuste. » La conséquence est qu'on ne peut attendre de la part des IA génératives un raisonnement enchaînant des inférences comme le ferait un démonstrateur de théorèmes en logique symbolique, ni un calcul fiable comme le ferait un système de calcul formel. L'IA recueille en parallèle des éléments en faisant fonctionner une multiplicité d'agents avec des profils distincts, et les exploite en simulant un jeu social de façon à fonder sa production. Significativement, MistralAI et ChatGPT résistent vigoureusement à une remise en cause de leurs conclusions sur le nombre de factorisations de $x^6 - 1$, indice de ce que les multiples agents convergent vers un faible nombre, l'idée même de pluralité des factorisations semblant leur poser problème.

Quand une stratégie de chaîne de pensées, c'est-à-dire un paramétrage demandant à un LLM de produire une suite d'énoncés conduisant à sa réponse, est mise en œuvre, cette suite n'est qu'en apparence une série d'inférences et reste loin d'une preuve mathématique, quoi qu'en disent certains auteurs. Un journaliste spécialisé, Baudart (2025), met par exemple en avant la supposée productivité d'une stratégie de chaîne de pensées pour un problème impliquant deux quantités inconnues non nulles Y et B et un nombre entier naturel n qu'il s'agit de déterminer. Un raisonnement mathématique montre que $Y = 16(6 - n)$ et $B = 24(n - 4)$. Il en résulte que n ne peut être égal qu'à 5. La chaîne de pensées produite par le LLM est bien différente. Le LLM fait divers calculs, puis émet une hypothèse sur la quantité Y qui, grâce aux calculs qui précèdent, conduit à la bonne valeur : *Toutes les conditions sont satisfaites si $n=5$. Donc, la réponse est 5.* On notera que, contrairement à ce que dit l'auteur la chaîne de pensées n'est pas un raisonnement, mais plutôt un choix au hasard qui, par chance, se révèle productif. De plus, dans un contexte mathématique, la résolution n'est pas achevée car la question de l'unicité n'est pas posée. Ajoutons que les chaînes de pensée, produites par les LLM sont très

⁸ The LLMs' high performance variance on different instances of the same question, their significant drop in performance with a slight increase in difficulty, and their sensitivity to inconsequential information indicate that their reasoning is fragile and may be more akin to sophisticated pattern matching rather than true logical reasoning (Mirzadeh et al 2025).

verbeuses. Si des enseignants sont tentés de les mettre à profit pour diriger les élèves vers une preuve construite, ils doivent être conscients du travail de lecture et de mise en cohérence que cela suppose chez les élèves. Jonin (2025) précise que selon lui, le « chain-of-thought prompting », ne permet pas de faire expliciter à l'IA un raisonnement dont il est incapable. En revanche, cet enseignant expert propose d'inciter les élèves à émettre des prompts ciblés, ce qui leur permet à la fois un usage efficace du LLM et la formulation de questions et de démarches structurées.

Transposé dans une situation où l'intelligence est celle d'humains, le processus de raisonnement des LLM est celui que Duval (1992) désigne par *argumentation*. De façon intéressante, l'IA générative reproduit avec 30 ans de retard l'évolution dans l'enseignement : dans les années 1970, l'accent dans l'enseignement est mis sur la démonstration comme enchaînement d'inférences, puis, devant le succès très relatif de cette approche, une bascule se produit vers des processus d'argumentation impliquant une dimension sociale. L'intelligence artificielle, quant à elle, a longtemps recouru à des systèmes à base de règles dans le but de créer des raisonnements mathématiquement valides, et, aussi proches que possible d'un raisonnement humain. L'introduction des LLM constitue une rupture en introduisant une multiplicité d'agents chacun mettant en œuvre de façon indépendante des structures créées par apprentissage. En mathématiques et dans leur l'enseignement, l'argumentation n'est pas la finalité : la preuve, élaborée selon les critères d'une communauté mathématique est ce qui fait foi. L'argumentation s'insère comme une phase ou un moment dans un processus où la réfutation joue un grand rôle. Ajoutons que curieusement les chaînes de pensée produites par les LLM ressemblent à ce que l'enseignement des math désigne par *narration de recherche*, une activité favorisant la rédaction, elle aussi préliminaire à une preuve mathématique.

L'interaction avec un LLM et le rôle des garde-fous,

Pour Storchan, « dans la plupart des applications, le modèle d'IA n'est pas utilisé de manière isolée, mais est intégré dans un échafaudage plus large comprenant divers composants, notamment des garde-fous destinés à contrôler les contenus générés, » L'objectif principal des garde-fous est de s'assurer que chaque réponse générée par le système d'IA est non seulement pertinente, mais aussi sûre, conforme aux réglementations en vigueur et alignée avec les valeurs et les objectifs de l'organisation. Les garde-fous peuvent aussi s'adapter au profil de l'utilisateur, en s'inspirant de ses prompts et de ses réactions aux réponses. La nécessité d'un garde-fou est évidente quand l'interaction avec l'IA porte sur des sujets sensibles et présente un risque pour l'utilisateur ou contrevient à des règles morales ou légales. Ces garde-fous reposent sur le même fonctionnement que l'IA elle-même (apprentissage par renforcement). Ils présentent donc la même opacité. Qu'en est-il lorsque l'interaction porte sur les mathématiques ? Et quand il s'agit d'un apprenant ?

L'expérience d'une interaction avec ChatGPT et MistralAI sur la factorisation montre des IA sûres d'elles, y compris quand le résultat est faux : factorisations visiblement fausses pour ChatGPT et rejet par MistralAI d'une factorisation exacte mais qui n'est pas dans la liste qu'il a donnée. Il faut de la persévérance de la part de l'utilisateur pour que l'IA admette qu'il y a une erreur. Elle présente alors des excuses obséquieuses. Mais sur les forums des utilisateurs mentionnent au contraire une tendance de l'IA à se plier à leur avis même s'il n'est pas

pertinent. L'IA peine donc à trouver un équilibre entre la confiance dans son propre raisonnement et l'ouverture aux rétroactions de l'utilisateur. Ceci doit être présent à l'esprit si l'on veut enclencher une dialectique preuve et réfutation avec une IA.

Que conclure ?

En dépit de la nouveauté du sujet, tant de choses se disent sur l'IA générative que je ressens une saturation de l'espace de discussion. Dans cet espace le risque est d'enfoncer des portes ouvertes ou de s'engager dans des perspectives romanesques. À côté de cela ce texte présente plusieurs particularités. En premier lieu il fait le lien avec un phénomène plus ancien, celui de l'irruption du calcul formel dans l'enseignement des maths. Qu'en est-il aujourd'hui de ce phénomène ? Il semble que du côté des praticiens dans beaucoup de pays l'attention se soit portée sur les examens avec des questions telles que : dans quelle mesure les examens permettent-ils d'utiliser le calcul formel et quel effet cela a-t-il sur les usages ? Du côté de la recherche, particulièrement dans un pays qui a autorisé l'usage du calcul formel, un certain désenchantement est noté : Jankvist et al (2019) basent leur analyse de l'usage du calcul formel dans un pays où les pratiques se limitent à ce que je qualifie de « techniques presse-bouton ». Ils font ainsi un constat qui ne peut conduire qu'à un retour en arrière. L'idée de tâches et de techniques nouvelles, génératrices de nouvelles praxéologies en interaction avec les techniques habituelles, développée dans les années 2000, s'est perdue alors que l'approche instrumentale prenait une direction psychologisante.

Une seconde particularité de ce texte est donc de réactiver ce cadre des techniques qui s'est trouvé marginalisé dans ce qu'est devenue l'approche instrumentale, et de le coordonner avec un autre cadre, les ETM, qui a démontré sa robustesse pour l'analyse des situations d'apprentissage. Les techniques sont étudiées sur des situations précises – la dialectique construction molle construction robuste avec la GD, et la factorisation des polynômes avec le calcul formel –, s'appuyant sur une analyse en profondeur du fonctionnement de ces artefacts dans le domaine sur lequel porte la situation. Le même type d'analyse est mené sur la factorisation avec l'IA générative. L'interaction avec deux LLM conduit à 3 constats : (1) Malgré leur plus grande souplesse, les LLM restent fixés sur une supposée unicité de la factorisation : la pseudo-transparence déjà mise en évidence avec le calcul formel persiste. (2) Au-delà de sa fluidité, le discours des LLM laisse supposer une influence d'algorithmes opaques. Il peut tantôt rester buté sur un résultat faux ou aussi bien se plier à une objection de l'utilisateur pas nécessairement fondée. (3) Ce même discours présente un défaut de cohérence plus ou moins prononcé, qui suggère que l'hétérogénéité des sources servant à l'entraînement de ces systèmes n'est pas maîtrisée.

Une étude du fonctionnement des LLM approfondit ces points. Les discours sur les IA génératives qui partent de limites constatées à un moment donné n'ont pas grand intérêt car ces limites peuvent être bientôt dépassées. En revanche, pour une situation d'enseignement donnée, il est utile d'explorer ces limites de façon à construire des situations didactiques où les élèves auraient à dépasser ces limites. Le processus de raisonnement des LLM relève de l'argumentation, c'est-à-dire de l'accumulation d'indices divers convergeant vers un résultat, plutôt que d'une preuve. En mathématiques et dans leur l'enseignement, l'argumentation s'insère comme une phase ou un moment dans un processus où la réfutation joue un grand

rôle. Les LLM quant à eux fonctionnent sur le mode de la majorité des avis, un fonctionnement avec lequel la preuve introduit une rupture. Passer de l'argumentation à la preuve reste un enjeu didactique. Il en découle que l'usage d'un LLM sur une question donnée n'a un intérêt dans une situation d'enseignement que si des techniques d'interrogation ont été instituées et si le travail débouche sur une exploitation des réponses du LLM pour une élaboration conforme aux standards institutionnels. J'ai noté aussi que l'IA peine à trouver un équilibre entre la confiance dans son propre raisonnement et l'ouverture aux rétroactions de l'utilisateur. Cela implique pour l'utilisateur de faire lui aussi preuve d'opiniâtreté et de coordonner avec d'autres instruments dont bien sûr le papier crayon.

J'illustre ces points par l'exploration des factorisations avec deux LLM. Partant de la mise en évidence d'une supposée unicité de la factorisation, il m'est apparu intéressant de questionner les LLM sur *toutes les factorisations de $x^6 - 1$* . La valeur $n=6$ a été choisie pour avoir un nombre significatif de diviseurs (4) entraînant un assez grand nombre de factorisations (14). Il a fallu rechercher des prompts adaptés, notamment en fournissant au LLM des factorisations qu'il ne trouvait pas. Il a fallu coordonner cette technique d'interrogation avec des techniques calcul formel et papier crayon, ainsi que l'usage de Wikipedia pour le nombre de partitions d'un ensemble à 4 éléments. La recherche de *toutes les factorisations dans \mathbf{Z}* pour une valeur de n bien choisie s'est trouvée être une tâche adéquate pour l'usage d'un LLM donné à un moment donné, dans le domaine des factorisations. Chaque nouvel instrument implique de nouvelles tâches pour de nouvelles techniques et de nouvelles conceptualisations. Il y a du travail pour la didactique et j'espère que les quelques éléments théoriques apportés dans ce texte y contribueront.

Pour une approche mathématique et humaine de l'IA

Je souhaite dans cette dernière section aller plus loin qu'une approche strictement didactique, bien qu'en apparence, mon propos ne soit pas directement relié à ce qui précède. Les mathématiques et leur enseignement ont une responsabilité sociale qui devient particulièrement visible avec les LLM. Je propose trois thèmes qui peuvent faire l'objet de travaux des élèves dans un cours de mathématiques.

La dimension mathématique des LLM.

En dépit de la complexité des différents modules et bases de données qui interviennent dans l'apprentissage et le fonctionnement des LLM, les élèves peuvent devenir conscients des principes de base des réseaux neuronaux et de l'apprentissage par renforcement. Je propose en annexe un exercice de niveau Première ou Terminale faisant travailler sur un réseau très simple la notion de fonction d'activation, de rétroaction et de descente de gradient. Cet exercice a été testé par un élève de Terminale qui s'en est très bien sorti en deux heures, ce temps comprenant l'apprentissage de fonctions du tableur intervenant dans l'exercice.

La dimension écologique des LLM.

Il est possible de faire travailler les élèves sur les données en consommation d'énergie de diverses « intelligences ». Des données sont disponibles sur <https://drane-versailles.region-academique-idf.fr/spip.php?article1167>. Citons la conclusion : *Le cerveau humain a une puissance brute de calcul 1000 fois plus grande que celle du meilleur processeur GPU actuel*

pour une puissance consommée 10 fois plus faible ! C'est l'occasion de bons exercices sur les ordres de grandeur.

La dimension sociétale des LLM.

Les études de l'impact de l'IA générative sur l'emploi convergent pour mettre en évidence une baisse relative de l'emploi des travailleurs en début de carrière, ce que seront nos élèves dans un futur proche, dans les métiers les plus exposés à l'IA. Il est à prévoir que l'amélioration des performances des LLM amplifie ce phénomène qui, pour certains, va prendre une dimension sociétale. Prenons le cas des métiers du développement logiciel qui prennent de multiples formes avec l'envolée des services numériques et des objets connectés et sont souvent assurés par des jeunes de formation scientifique en début de carrière. Il est vrai que l'usage d'un LLM permet un gain de temps considérable sur la recherche d'un traitement adapté et son codage dans un langage et sur une machine spécifique, ce qui ne veut pas dire que le LLM fait le travail tout seul. L'activité de développement s'oriente vers la production de spécifications précises et la vérification soignée du code produit par un LLM. Elle demande une moindre spécialisation relative à un langage ou à un environnement, mais aussi la capacité à en saisir les principes : pour programmer par exemple un grille-pain connecté, il faut être capable d'émettre une spécification adaptée à des utilisateurs et des ustensiles génériques, et aussi avoir une connaissance suffisante du langage de commande et du fonctionnement interne de l'ustensile, faute de quoi prévoir un extincteur ! Une même analyse peut être faite des métiers d'analyste de données. Dans ce basculement des métiers vers une plus grande autonomie vis-à-vis de procédures spécifiques et une capacité à s'approprier des environnements divers, l'enseignement des mathématiques, en promouvant des techniques d'usage de LLM, a un grand rôle à jouer.

Références

- Aghion, P., Bouverot, A. (2025) *I.A. Notre ambition pour la France*. Paris : Odile Jacob
- Artigue, M. (1997). Le logiciel Dérive comme révélateur de phénomènes didactiques liés à l'utilisation d'environnements informatiques pour l'apprentissage. *Educational Studies in Mathematics*, 33(2), 133-169
- Artigue, M. (2002). Learning mathematics in a CAS environment: The genesis of a reflection about instrumentation and the dialectics between technical and conceptual work. *International Journal of Computers for Mathematical Learning*, v7 n3 p245-274.
- Baudard, G. (2025) Faire communiquer un modèle de langage et un assistant de preuve. <https://www.lemonde.fr/blog/binaire/2025/05/23/faire-communiquer-un-modele-de-langage-et-un-assistant-de-preuve/>
- Chevallard, Y. (1999) L'analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en didactique des mathématiques*, 19(2), 221-266.
- Durand, J. (2025) Quelques expériences d'usage d'une IA en classe <http://revue.sesamath.net/spip.php?article1709>
- Duval, R. (1992) Argumenter, démontrer, expliquer : continuité ou rupture cognitive ? *Petit x*, n°31, 37-61.

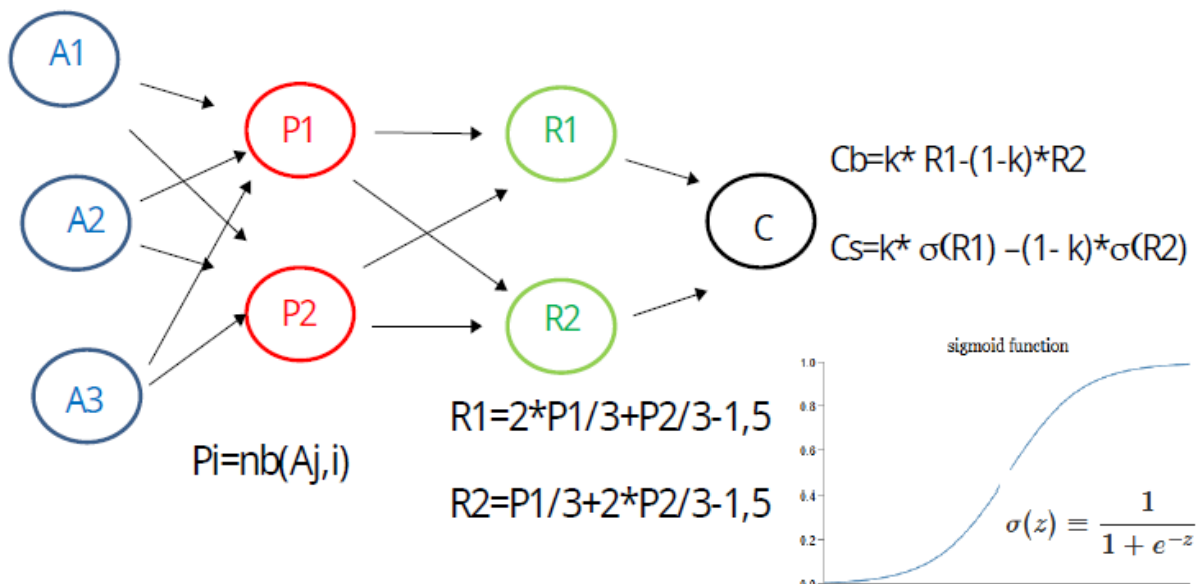
- Duval, R. (1995), Why to teach geometry, *Icmi Studies on Geometry*, Catania.
- Elbaz-Vincent, P. (2005). A CAS as an assistant to reasoned instrumentation. In D. Guin, K. Ruthven, et L. Trouche (Eds.), *The didactical challenge of symbolic calculators: turning a computational device into a mathematical instrument* (pp. 41–65). New York: Springer. 19
- Jankvist, U. T., Misfeldt, M. et Aguilar, M. S. (2019). What happens when CAS-procedures are objectified? – the case of “solve” and “desolve”. *Educational Studies in Mathematics*, 101(1), 67-81.
- Jonin, P. (2025) Utiliser l’IA générative en classe de mathématiques et en SNT <http://revue.sesamath.net/spip.php?article1691>
- Kuzniak, A. (2013). Teaching and learning geometry and beyond... In B. Ubuz, Ç. Haser, et M. A. Mariotti (Eds.), *Proceedings of CERME 8* (pp. 33–49). Antalya, Turkey.
- Kuzniak, A., Tanguay, D. et Elia, I. (2016). Mathematical Working Spaces in schooling: an introduction. *ZDM Mathematics Education* 48, 721–737.
- Kuzniak, A. (2022). The Theory of Mathematical Working Spaces—Theoretical Characteristics. In: Kuzniak, A., Montoya-Delgado, E., Richard, P. R. (eds) *Mathematical Work in Educational Context*. Springer, Cham. https://doi.org/10.1007/978-3-030-90850-8_1
- Lagrange J.-B., Artigue M., Laborde C., Trouche L. (2003) Technology and Mathematics Education: A Multidimensional Study of the Evolution of Research and Innovation. In: Bishop A.J., Clements M.A., Keitel C., Kilpatrick J., Leung F.K.S. (eds) *Second International Handbook of Mathematics Education*. Springer, Dordrecht.
- Lagrange, J.-B. (2000). L'Intégration d'Instruments Informatiques dans l'Enseignement : Une Approche par les Techniques (The Integration of technological Instruments in teaching/learning: An Approach from the techniques). *Educational Studies in Mathematics*, 43(1), 1-30.
- Lagrange, J.-B. (2005). Transposing computer tools from the mathematical sciences into Teaching. In D. Guin, K. Ruthven, et L. Trouche (Eds.), *The didactical challenge of symbolic calculators* (pp. 67-83). New York: Springer.
- Lagrange, J.-B., Richard, P. R., Vélez, M.P., Van Vaerenbergh, S. (2023). Artificial Intelligence Techniques in Software Design for Mathematics Education. In: Pepin, B., Gueudet, G., Choppin, J. (eds) *Handbook of Digital Resources in Mathematics Education*. Springer, Cham. https://doi.org/10.1007/978-3-030-95060-6_37-1
- Lagrange, J.-B., Richard, P.R., (2022). Instrumental Genesis in the Theory of MWS: Insight from Didactic Research on Digital Artifacts. In: Kuzniak, A., Montoya-Delgado, E., Richard, P. R. (eds) *Mathematical Work in Educational Context*, Springer International Publishing, pp.211-228,, [10.1007/978-3-030-90850-8_9](https://doi.org/10.1007/978-3-030-90850-8_9). hal-03636678
- Lesnes-Cuisiniez, E. (2021). *Modélisation didactique de parcours d'apprentissage dans un EIAH pour l'entrée dans le raisonnement géométrique au cycle 4, en appui sur les*

problèmes de construction de figures planes. Thèse. Université de Paris, Français. NNT : tel-03404996v1.

- Mirzadeh, I., Alizadeh, K., Shahrokhi, B. et al (2025) Gsm-symbolic: understanding the limitations of mathematical reasoning in large language models. <https://doi.org/10.48550/arXiv.2410.05229>
- Mounier G., Aldon G. (1996) A Problem Story: Factorisations of $x^n - 1$. *International DERIVE Journal*, vol.3. 3, pp. 51-61
- Pepin, B., Buchholtz, N. et Salinas-Hernández, U. (2025) A Scoping Survey of ChatGPT in Mathematics Education. *Digit Exp Math Educ* 11, 9–41. <https://doi.org/10.1007/s40751-025-00172-1>
- Richard, P. R., Venant ,F., Gagnon ,M. (2019) Issues and Challenges in Instrumental Proof. In: Hanna G., Reid D., de Villiers M. (eds) *Proof Technology in Mathematics Research and Teaching*. Springer, Cham.
- Trouche, L., (1994) Calculatrices graphiques : la grande illusion. *Repères IREM*, 14, pp.39-55. {hal-01541174}
- Vérillon, P. et Rabardel, P. (1995). Cognition and artefacts: A contribution to the study of thought in relation to instrumental activity. *European Journal of Psychology of Education*, 10, 77-101.

Annexe. Un exercice pour travailler sur les réseaux neuronaux

On considère un réseau neuronal devant permettre à trois amis A_1, A_2, A_3 de choisir parmi deux restaurants r_1 et r_2 . Le but est de comparer deux fonctions d'activation, C_b (binaire) et C_σ (sigmoïde). Comme on le verra C_σ est calculé à partir de la fonction sigmoïde (ou logistique) définie sur les réels par $\sigma(z) \equiv \frac{1}{1 + e^{-z}}$. L'exercice se poursuit par un exemple d'apprentissage par rétroaction.



Le choix des amis est r_1 si la fonction d'activation est strictement positive et r_2 si la fonction est strictement négative. Une couche intermédiaire comprend deux nœuds P_1 et P_2 correspondant à deux plats favoris des amis. P_1 envoie le nombre n_1 des amis qui apprécient le premier plat et P_2 le nombre n_2 des amis qui apprécient le second (appréciation non exclusive).

- Déterminer les valeurs possibles des couples (n_1, n_2) . Entrez ces valeurs dans deux colonnes P_1 et P_2 d'un tableur.
- r_1 fait mieux le premier plat. r_2 fait mieux le second plat. Le nœud R_1 envoie donc la valeur $\frac{2n_1}{3} + \frac{n_2}{3} - \frac{3}{2}$ et le nœud R_2 envoie la valeur $\frac{n_1}{3} + \frac{2n_2}{3} - \frac{3}{2}$. ($2/3$ et $1/3$ sont des poids, et $-3/2$ assure une moyenne nulle ce qui permettra une prise de décision en fonction du signe). Dans les deux colonnes suivantes faire calculer les valeurs sortant des nœuds R_1 et R_2 .
- On considère qu'il existe d'autres différences influant le choix des amis (situation, accueil...) On introduit un poids k (compris entre 0 et 1) pour quantifier cette influence.
Choisir une case du tableur pour le poids k . Dans la colonne suivante calculer la fonction de choix $C_b = kR_1 - (1 - k)R_2$ puis la colonne correspondant au nœud final (r_1 si $C_b > 0$; r_2 si $C_b < 0$; 0 si $C_b = 0$).
- Noter et représenter graphiquement le nombre de choix 0, r_1 et r_2 pour chaque valeur de k de 0 à 1 par pas de 0,1. Vous pouvez le faire avec le tableur en créant un curseur (slider) pour faire varier k , ou avec un programme Python ou encore Geogebra. Que remarque-t-on ? En l'absence de la connaissance de (n_1, n_2) , ces nombres de choix permettent-ils une prise de décision ?

5. Dans cette question et les suivantes, on considère la fonction d'activation $C_\sigma = k\sigma(R_1) - (1 - k)\sigma(R_2)$. Remplacer C_b par C_σ dans le tableur. Noter et représenter de même le nombre de choix $(0, r_1, r_2)$ pour les mêmes valeurs de k . Comparer avec les résultats de la question 3. Commentez l'assertion suivante : « la fonction d'activation sigmoïde permet qu'une petite modification sur un poids entraîne une petite modification sur le résultat ».
6. Vous avez observé que pour $k=0,5$, il y a 6 choix r_1 et 6 choix r_2 . Une étude statistique montre que le restaurant 2 n'est choisi que dans 5 cas (n_1, n_2) . En faisant varier k (vous pouvez changer l'incrément), repérer qu'il existe un couple (n_1, n_2) qui fait basculer le choix correspondant de r_2 de 6 à 5 ainsi qu'un couple (n'_1, n'_2) qui fait basculer le choix correspondant de r_2 de 5 à 4. Cette observation va permettre d'estimer le poids k .
7. Exprimer la fonction de choix $C_\sigma(k)$ pour le couple (n_1, n_2) . Quelle équation faut-il résoudre pour trouver la valeur de k pour laquelle le nombre de choix r_2 bascule de 6 à 5 ? Résoudre.
8. Calculer de même la valeur de k pour le couple (n'_1, n'_2) qui fait basculer le nombre de choix r_2 de 5 à 4. Vous obtenez ainsi un encadrement de la valeur de k correspondant à l'observation statistique.
9. Vous avez résolu les équations des questions 6 et 7 par l'algèbre. Dans la pratique des réseaux neuronaux, ces équations impliquent de très nombreuses inconnues (tous les poids du réseau) et ne sont pas résolubles par l'algèbre. L'actualisation des poids à partir d'exemples (l'apprentissage) utilise un algorithme de "descente de gradient". Nous écrivons l'équation trouvée à la question 6 sous forme $f(k) = 0$. L'algorithme part de la valeur $k = 0.5$, pour laquelle $f(k)$ est environ égal à $-0,03$. Il ajoute itérativement une valeur de plus en plus petite jusqu'à obtenir $f(k) > 0$. Par exemple, cette valeur peut être choisie égale à $-a \cdot f(k)$, a étant un facteur à adapter. Implémenter cet algorithme en python ou sur le tableur avec $a = 1/3$ Comparer au résultat obtenu en 6. Que se passe-t-il si on prend un facteur plus grand (par exemple $a = 1$) ou plus petit (par exemple $a = 1/5$) ? Rechercher dans une documentation sur les réseaux neuronaux le nom donné à $f(k)$ et celui donné au facteur a .