L'IA veut-elle prendre nos emplois de mathématiciens en recherche?

Articles contribués

Octobre 2026 (tome 58, no. 5)

George A. Grätzer

Résumé. ChatGPT, accessible au grand public, peut-il contribuer à la recherche en mathématiques? Je décris une expérience au cours de laquelle je lui ai demandé de trouver une démonstration plus courte d'un théorème portant sur les ensembles ordonnés finis. Je venais tout juste de trouver moi-même une telle démonstration.

À partir des articles pertinents et d'une brève consigne, ChatGPT a produit une démonstration courte en environ cinq minutes, en adoptant une approche complètement différente de la mienne. Cette expérience a rendu la question du titre particulièrement personnelle.

D'abord, ils sont venus chercher les socialistes, et je n'ai rien dit, parce que je n'étais pas socialiste.
Ensuite, ils sont venus chercher les syndicalistes, et je n'ai rien dit, parce que je n'étais pas syndicaliste.
Ensuite, ils sont venus chercher les Juifs, et je n'ai rien dit, parce que je n'étais pas juif.
Ensuite, ils sont venus me chercher, et il ne restait plus personne pour prendre ma défense.

Martin Niemöller

Luddites du monde entier, unissez-vous. D'après Karl Marx

1. Introduction

Dans deux éditoriaux récents, Robert Dawson [1, 2] présente un récit intéressant de ce qui s'est passé lorsqu'OpenAI a décidé de dépenser des millions de dollars pour résoudre toute une série de vieux problèmes mathématiques. Cette histoire m'a particulièrement interpellé.

Laissez-moi vous expliquer. En 1963, le résultat suivant a été publié [12] : soit L un treillis algébrique. Il existe alors une algèbre A telle que le treillis des congruences de A soit isomorphe à L.

Peu après, ce résultat est devenu connu sous le nom de théorème de Grätzer-Schmidt. À peu près à la même époque, l'expression « théorème de Grätzer-Schmidt fini » a été introduite. Il ne s'agissait toutefois pas d'un théorème, mais d'un problème ouvert : si L est fini, peut-on choisir A de façon qu'elle soit finie?

Classification des sujets mathématiques de 2020. Sujet principal : 06A07; sujet secondaire : 06B05. Mots et expressions clés : intelligence artificielle, ChatGPT, recherche mathématique, ensembles ordonnés, applications qui préservent l'ordre.

Et les millions de dollars dépensés ont permis de répondre par la négative à cette question.

Cependant, cet article ne porte pas sur les millions de dollars qu'OpenAI a dépensés pour résoudre de vieux problèmes mathématiques. Il porte sur le ChatGPT auquel vous et moi avons accès.

2. Service à la clientèle

La plupart d'entre nous avons d'abord fait l'expérience du remplacement du personnel du service à la clientèle par l'IA. Dans la plupart des cas, les résultats n'étaient pas très bons, et il fallait crier « Je veux parler à un vrai agent! » à maintes reprises avant d'être mis en contact avec quelqu'un capable de répondre à nos questions. Mais les choses sont en train de changer.

Il y a quelques semaines, j'avais très hâte d'essayer certaines des nouvelles fonctionnalités de macOS 27. J'ai donc installé une version bêta. Malheureusement, elle comportait tellement de bogues que j'ai dû revenir à macOS 26. Ce n'est pas une opération facile.

Sur un Mac, il est assez simple de mettre à niveau le système d'exploitation, mais revenir à une version antérieure est beaucoup plus compliqué. J'ai appelé Apple et j'ai été pris en charge par le service à la clientèle utilisant l'IA, qui m'a guidé tout au long de cette procédure difficile et complexe, avec un succès total.

D'abord, ils sont venus chercher les spécialistes du service à la clientèle, et je n'ai rien dit, parce que je n'étais pas spécialiste du service à la clientèle.

3. Programmeurs et programmeuses

Il y a quelques mois, plusieurs articles de journaux expliquaient pourquoi les personnes récemment diplômées en informatique n'arrivaient pas à trouver d'emploi comme programmeurs et programmeuses : l'IA programme mieux qu'eux.

Ensuite, ils sont venus chercher les programmeurs et programmeuses, et je n'ai rien dit, parce que je ne suis pas programmeur.

4. Mathématiciennes et mathématiciens en recherche

Après avoir téléversé un article sur arXiv, j'ai reçu un avertissement m'informant que je ne pouvais soumettre que deux articles par mois, car il y avait trop de soumissions. Nous avons tous entendu parler des revues mathématiques qui croulent sous les articles rédigés par l'IA.

Mais comment mesurer l'efficacité de l'IA, accessible à vous et à moi, dans la rédaction d'articles mathématiques? J'ai eu une occasion unique de mettre cette capacité à l'épreuve.

Permettez-moi de poser le contexte. Pour les ensembles ordonnés A et B, notons AB l'ensemble ordonné de toutes les applications de B dans A qui préservent l'ordre, ordonné composante par composante.

En 1978, Dwight Duffus a proposé le problème suivant dans sa thèse de doctorat [3] : soient A et B deux ensembles ordonnés finis non vides. Si AA est isomorphe à BB, peut-on en conclure que A est isomorphe à B? Autrement dit, peut-on retrouver A à partir de AA?

Les références comprennent des publications sur ce sujet datant de la fin des années 1970 et des années 1980. J. D. Farley a ravivé l'intérêt pour ce problème au début des années 2020. Celui-ci a récemment été résolu par J. D. Farley et G. Grätzer [9]. La démonstration est assez longue et fait appel à des résultats approfondis de J. Hashimoto [13], L. Lovász [14] et R. McKenzie [15].

Il était donc logique de chercher une démonstration plus courte. J'en ai trouvé une; voir les sections 2 et 3 de G. Grätzer [10]. Cela m'a donné une occasion unique de vérifier si ChatGPT était aussi doué que moi pour trouver une démonstration courte.

Il convient de noter que je suis particulièrement bien placé pour trouver une démonstration courte, compte tenu de mon expérience en théorie des treillis; voir G. Grätzer [11].

J'ai téléversé dans ChatGPT tous les articles pertinents cités dans les références, accompagnés d'une brève consigne : « Trouve une démonstration courte. » En environ cinq minutes, ChatGPT m'a répondu en proposant une démonstration courte; voir la section 4 de G. Grätzer [10].

L'approche de ChatGPT est complètement différente de la mienne. J'ai donc trouvé une démonstration courte et ChatGPT en a trouvé une autre, et les deux ont été publiées dans [10].

Ensuite, ils sont venus chercher les mathématiciennes et les mathématiciens, et il ne restait plus personne pour prendre ma défense.

5. Présentation des trois démonstrations

5.1. L'idée commune aux trois démonstrations. La hauteur de X est la longueur d'une chaîne de longueur maximale. Le long d'une chaîne dans XY, chacune des |Y| valeurs peut augmenter au plus (X) fois. Comme il est facile de construire une telle chaîne, on obtient, hauteur(XY)=|Y|hauteur(X). On dit qu'un intervalle [a,b] de X est de hauteur maximale si sa hauteur est égale à hauteur(X). On note MaxInt(X) l'ensemble de ces intervalles. Les intervalles sont bornés, et les ensembles ordonnés bornés sont précisément ceux dont les puissances sont tempérées. Soient Y1, . . . , Yr les composantes de Y.Si [f, g] est un intervalle de hauteur maximale de XY, le dénombrement ci-dessus force chaque [f (y), g(y)] à être de hauteur maximale dans X, de sorte que f(y) est minimal et g(y) est maximal dans X; les éléments minimaux comparables coïncident, de sorte que f et g sont constants sur chaque composante, et g] ∼= IY1 1 × · · · × IYrr , Ii ∈ MaxInt(X). Réciproquement, chaque choix des Ii se présente, et la correspondance est une bijection. Pour A connexe, cela signifie que les intervalles de hauteur maximale deE = AA sont les puissances LA avec L ∈ MaxInt(A) : l'exposant A réapparaît, rattaché à une base bornée. Deux difficultés subsistent, et les preuves divergent exactement ici. (i) Il existe de nombreux intervalles de hauteur maximale ; pour A non connexe, ce sont des produits mixtes. Quel objet bien défini E produit-il ? (ii) On a besoin d'un logarithme tel que(LA) ∼= A × L(L) pour L borné, puis d'un moyen de diviser par l'inconnu L(L).

5.2. La preuve longue. Logarithme. Farley et Grätzer le construisent à partir de la théorie de la structure. Pour P borné, toute application dans PY est reliée à une application constante, de sorte que PY est égal au sous-ordre Con(P, Y)de composantes contenant des applications constantes sur chaque composante de Y. La forme normale de McKenzie écrit tout ensemble connexe directement indécomposable D de manière unique sous la forme D ∼= Con(ED, FD avec ED absolument indécomposable et FD connexe ; une puissance par un ensemble connexe Y conserve ED et remplace FD par FD × Y .. Avec la factorisation de Hashimoto P ∼= D1 × · · · × Dk, posons Tow (P) = FD1 + · · · + FDk . Alors Tow(P Q) ∼= Tow P + Tow Q, Tow(P Y) ∼= Y × Tow P (P Borné), et height(Tow P) < height P pour P non trivial. C'est un véritable logarithme sur tous les ensembles ordonnés connexes, qui repose sur deux théorèmes profonds. Un seul intervalle. Pour (i), la preuve longue sélectionne un unique intervalle. Elle construit, en trois étapes lexicographiques (vecteurs d'exposants des factorisations de Hashimoto, puis paires de formes normales, puis fonctions de multiplicité), un ordre linéaire $\preceq$ des types d'isomorphisme d'ensembles ordonnés connexes tel que P < Q ⇒ P × R < Q × R et P Y < QY . Alors Max(X), the ≤-le plus grand élément pour MaxInt(X), st bien défini et satisfait Max(XY) ∼= Max X × Max Y et Max(XY) ∼= Max(X)Y . Reconstruction. Avec L = Max A et R = Tow L, les règles donnent, à partir de E seul, U = Tow(Max E) ∼= R × A, V = Max U ∼= Max R × L, W = Max(Tow V) ∼= Max R;la dernière étape utilise la chute de hauteur pour déterminer quel terme de la somme Tow V ∼= Tow(Max R) + R porte les intervalles supérieurs. Simplifier par W dans V donne L,d'où R;simplifier par R dans U donne A. La simplification repose à nouveau sur Hashimoto : on code un ensemble ordonné sous forme d'un polynôme en des indéterminées indexées par les indécomposables et on simplifie dans l'anneau de polynômes. L'itération des mêmes opérations permet de traiter des tours telles que (AA)A et A(AA), ainsi que tout parenthésage d'au plus quatre copies de A.

5.3. Les remplacements élémentaires. Les deux preuves courtes conservent le lemme des intervalles et écartent le reste. Logarithme à partir des couvertures inférieures. Soit J(B) l'ensemble des éléments ayant exactement une couverture inférieure. Pour B borné, les éléments de BY ayant une unique couverture inférieure sont les applications égales à un certain j ∈ J(B) sur une section finissante principale ↑ y et à oB ailleurs : abaisser une application en un point minimal de son support produit une couverture inférieure, de sorte que l'unicité fixe un point, une valeur et aucune variation au-dessus. D'où J(BY) ∼= J(B) × Y ∂; et puisqu'une couverture dans un produit déplace une coordonnée le long d'une couverture, J(BC) ∼= JB + JC. Avec L(B) = J(B)∂, c'est exactement le comportement en trois lignes de Tow, maintenant pour B borné uniquement, y compris la chute de hauteur, et cela prend une page à démontrer. Tous les intervalles au lieu d'un seul. Le lemme des intervalles est une bijection MaxInt(XY) ↔ MaxInt(X)r, et les intervalles supérieurs de XY sont les produits I × J. Toute fonction symétrique de tous les intervalles hérite donc des règles de produit et de puissance, sans ordonnancement des types d'isomorphisme. Simplification par dénombrement. Pour Z connexe, soit hom(Z, X) le nombre d'applications croissantes Z → X ; il est additif sur les sommes, multiplicatif sur les produits et strictement positif. Lovász (1967) : Si hom(Z, X) = hom(Z, Y), pour tout Z connexe, alors X ∼= Y; le cas des ensembles ordonnés admet une courte preuve par inclusion-exclusion. Ainsi, XT ∼= Y T implique X ∼= Y : on applique hom(Z, –) et on divise par hom(Z, T) > 0.

5.4. Première preuve courte : prise de moyenne. Soit R l'ensemble des combinaisons rationnelles positives ou nulles d'ensembles ordonnés connexes, multipliées par le produit direct, un ensemble ordonné étant la somme de ses composantes. Pour X de hauteur strictement positive, posons Av(X) = 1 | MaxInt X| X I∈MaxInt X L(I),et prolongeons à R+ en prenant la moyenne sur les termes de hauteur maximale. Le lemme des intervalles et les règles sur L donnent Av(RS) = Av R+Av S et Av(XY) = Y ·Av X, et la chute de hauteur donne height(Av R) < height R. Avec R = Av A et S = Av B, on a A · R = Av(AA) = Av(BB) = B · S, d'où R + Av R = S + Av S. Puisque Av R se situe strictement en dessous de la hauteur maximale de R R, la partie supérieure de R + Av R est celle de R ; en prendre la moyenne redonne Av R, et soustraire redonne R. Ainsi R = S, et simplifier par R dans AR = BR par dénombrement donne A ∼= B. Le logarithme inconnu n'est jamais calculé ; il s'identifie par sa chute de hauteur et est éliminé par division.

5.5. Deuxième preuve courte : trois invariants. Pour s'en tenir aux entiers, soit M(X) la somme disjointe d'une copie de chaque intervalle de hauteur maximale de X ; alors M(XY) ∼= MX MY et M(XY) ∼= M(X)Y . Posons P = MA avec m composantes, et W = M(J P), avec k composantes. À partir de E = AA seul, formons U = J(M E), V = M U, Q = M(J* V), où J* compte les couvertures supérieures. En suivant J à travers une puissance à base non connexe, on obtient U ∼= µ(JP × A∂), V ∼= µ(W × P ∂), Q ∼= µk W ∂ , µ = mc–1, c étant le nombre de composantes

de A ; la chute de hauteur isole à nouveau le terme de la somme J* V portant les intervalles supérieurs, et A = height U − height Q. Dénombrons maintenant avec un ensemble connexe fixé Z, une fois pour A et une fois pour B. Les dénombrements pour Q et V se divisent pour donner kB hom(Z∂ , PA) = kA hom(Z∂ , PB), d'où kB PA ∼= kAPB et kB JPA ∼= kAJPB. En insérant ceci dans le dénombrement pour U, en simplifiant par des nombres strictement positifs, et en utilisant |A| = |B| (hauteurs égales et |A| height A = |B| height B) il reste hom(Z, A∂) = hom(Z, B∂) pour tout Z connexe. Le théorème de Lovász conclut.

References

[1] R. Dawson, Collaboration With Our New Robot Overlords. CMS Notes June 2026 (Vol. 58, No. 3).

[2] R. Dawson, Collaboration With Our New Robot Overlords. II. CMS Notes September 2026 (Vol. 58, No. 4).

[3] D. A. Duffus, Toward a Theory of Finite Partially Ordered Sets. Ph.D. thesis, University of Calgary, 1978.

[4] D. Duffus, Powers of ordered sets. Order 1 (1984), 83–92.

[5] D. Duffus and R. Wille, A theorem on partially ordered sets of order-preserving mappings. Proc. Amer. Math. Soc. 76 (1979), 14–16.

[6] J. D. Farley, Poset exponentiation and a counterexample Birkhoff said in 1942 he did not have. Order 39 (2022), 243–250.

[7] , Does the endomorphism poset P P determine whether a finite poset P is connected? An issue Duffus raised in 1978. Math. Bohem. 148 (2023), 435–446.

[8] , Does the endomorphism ordered set of a finite ordered set determine the ordered set? The cases of height 1 and 'trebled' ordered sets. ArXiv 2609.38532.

[9] J. D. Farley and G. Grätzer, The AA problem for finite ordered sets. Submitted for publication. Research Gate.

[10] G. Grätzer, Two Short Proofs of the AA Theorem. ResearchGate.

[11] G. Grätzer, Lattice theory: foundation, Birkhäuser/Springer Basel AG, Basel; 2011; pp. xxx+613.

[12] G. Grätzer and E. T. Schmidt, Characterizations of congruence lattices of abstract algebras, Acta Sci. Math. (Szeged); vol. 24; 1963; pp. 34–59.

[13] J. Hashimoto, On direct product decomposition of partially ordered sets. Ann. of Math. 54 (1951), 315–318.

[14] L. Lovász, Operations with structures. Acta Math. Acad. Sci. Hungar. 18 (1967), 321–328.

[15] R. McKenzie, Arithmetic of finite ordered sets: Cancellation of exponents, II. Order 17 (2000), 309–332.

Droits d'auteurs & autorisations

La Société mathématique du Canada autorise les lecteurs individuels de cette publication à copier les articles pour leur usage personnel. L'utilisation à d'autres fins est strictement interdite. Pour obtenir une licence autre que la copie d'articles à des fins personnelles, veuillez contacter la Société mathématique du Canada pour demander des autorisations ou des conditions de licence.

Société mathématique du Canada — 616 Cooper St., Ottawa (ON) K1R 5J2, Canada