Systèmes génératifs et résolution de problèmes mathématiques : La collection de cailloux

(version du 17.02.25)

Réussite étonnante dire qu'iel n'y comprend rien !

Anonyme (XXIe siècle)

Introduction

Ce compte rendu concerne, après le problème des dromadaires et des chameaux[1], un autre problème pour un travail exploratoire. Les premières tentatives datent de janvier 2023. Elles sont complétées pour comparaison par une demande qui date du début de la rédaction de ce compte rendu (fin décembre 2024).

Deux problèmes à propos de cailloux

Premier problème : Collection de cailloux (cat 3-4)

Le problème qui va être soumis à différents systèmes conversationnels est tiré de la banque de problèmes du Rallye mathématique transalpin (RMT). Il s'agit du problème 2 de la première épreuve du 29e RMT (2022)[2]. Il est de catégorie 3-4 (enfants de 9-10 ans).

> Jacques a ramassé 45 cailloux colorés et il veut les répartir en quatre boîtes. Il a mis des cailloux dans la première boîte. Dans la deuxième boîte, il a mis plus de cailloux que dans la première, mais moins que dans la troisième. Dans la troisième boîte, il a mis le double du nombre de cailloux contenus dans la première. Dans la quatrième boîte, il a mis le triple du nombre de cailloux contenus dans la première boîte.
>
> Combien de cailloux y a-t-il dans chaque boîte ?
>
> Montre comment tu as trouvé le nombre de cailloux dans chaque boîte.

Encadré 1 : Enoncé original

Une solution humaine

On note x et y le nombre de cailloux dans les deux premières boîtes. On a donc les contraintes :

$x+y+2x+3x=45$ et $x<y<2x$

$y=45-6x$ et donc $x<45-6x<2x$ on en déduit :

1) $7x<45$ d'où $x\leq 6$

2) $45<8x$ d'où $x\geq 6$

1 http://www.projet-ermitage.org/thema2/Chronique-chameaux.pdf

2 http://www.projet-ermitage.org/ARMT/navi_fic2.php?code=op143-fr&flag=1&langue=fr&enonce=29rmti_fr-2&w=0

En définitive les contenus des boîtes sont respectivement 6, 9, 12, 18 dont la somme vaut bien 45.

Autre méthode : considérer la relation $x=(45-y)/6$ et parcourir les valeurs de y possibles.

D'autres méthodes, par essais plus ou moins judicieux, peuvent être utilisées par des enfants ne disposant pas de l'outillage algébrique.

Les solutions des IA génératives

ChatGPT

Le premier essai (encadré 2) a été effectuée peu après la mise à disposition au grand-public de ChatGPT. Le passage à la formulation algébrique nous a, à l'époque, assez « bluffés » jusqu'à la relation $x = (45-y)/6$. Puis aucune méthode directe ne convient et le système s'en tient à reformuler le problème.

On peut résoudre ce problème en utilisant des équations à base de nombres incertains, et en utilisant des propriétés de l'addition et de la multiplication pour trouver des relations entre les différents nombres de cailloux dans chaque boîte.

1. On sait que Jacques a mis des cailloux dans la première boîte. On peut donc définir x comme étant le nombre de cailloux dans la première boîte.
2. On sait que dans la deuxième boîte, il y a plus de cailloux que dans la première, mais moins que dans la troisième. On peut donc définir y comme étant le nombre de cailloux dans la deuxième boîte, où x < y < z.
3. On sait que dans la troisième boîte, il y a le double du nombre de cailloux contenus dans la première. On peut donc écrire z = 2x comme étant le nombre de cailloux dans la troisième boîte.
4. On sait que dans la quatrième boîte, il y a le triple du nombre de cailloux contenus dans la première boîte. On peut donc écrire t = 3x comme étant le nombre de cailloux dans la quatrième boîte.

On sait que Jacques a ramassé 45 cailloux colorés, donc x + y + z + t = 45. On peut maintenant utiliser les relations établies entre x, y, z et t pour résoudre pour x.

x + y + 2x + 3x = 45

6x + y = 45

x = (45 - y)/6

On ne peut pas trouver une valeur numérique pour x, y, z ou t car on ne connait pas la valeur de y. Il faut donc utiliser l'énoncé pour déduire les nombres de cailloux dans chaque boîte. On peut voir qu'on a x < y < z < t On peut donc dire qu'il y a moins de cailloux dans la première boîte que dans la deuxième, moins de cailloux dans la deuxième boîte que dans la troisième, moins de cailloux dans la troisième boîte que dans la quatrième.

Encadré 2 : Solution de ChatGPT (version 3, janvier 2023)

L'idée qui rompt la séquence de déductions algébriques consiste à choisir y pour que $45 - y$ soit divisible par 6. Elle n'apparaît pas dans la solution délivrée par ChatGPT (y pourrait valoir 3, 9, 15 ... mais seul 9 convient pour que $y < z$). A noter que sur 766 classes d'enfants de 10 ans, 50% d'entre-elles parviennent à la solution et 30% en ont compris la logique.

Le deuxième essai (encadré 3) a eu lieu une année plus tard. La démarche adoptée est un peu différente. Est-ce dû à l'aspect probabiliste de la génération ou à une modification du système liée au passage de la version 3 à 3.5 ? mystère ! Plutôt que de tenter de résoudre algébriquement l'équation, le système *dit* procéder à des tentatives. Dans ce cas, la deuxième inconnue *y* est le surplus de la deuxième boîte par rapport à la première. La condition « contenu de la deuxième boîte inférieur à celui de la troisième » n'est pas respectée. Contrairement à l'annonce, il n'est pas clair si plusieurs valeurs différentes ont été essayées ?

Appelons x le nombre de cailloux dans la première boîte.

1. Dans la deuxième boîte, Jacques a mis plus de cailloux que dans la première, mais moins que dans la troisième. Donc, le nombre de cailloux dans la deuxième boîte est x + y, où y est un nombre positif.
2. Dans la troisième boîte, Jacques a mis le double du nombre de cailloux contenus dans la première. Donc, le nombre de cailloux dans la troisième boîte est 2x.
3. Dans la quatrième boîte, Jacques a mis le triple du nombre de cailloux contenus dans la première boîte. Donc, le nombre de cailloux dans la quatrième boîte est 3x.

La somme totale des cailloux dans les quatre boîtes est égale à 45, car Jacques a ramassé 45 cailloux au total.

Donc, l'équation est :

x + (x + y) + 2x + 3x = 45

Regroupons les termes similaires :

7x + y = 45

Maintenant, nous devons trouver des valeurs de x et y qui satisfont cette équation tout en respectant les conditions données.

Puisque Jacques a mis plus de cailloux dans la deuxième boîte que dans la première, mais moins que dans la troisième, cela signifie que y doit être strictement positif.

Essayons différentes valeurs pour x et y :

- Si x = 5 et y = 10, alors 7x + y = 7(5) + 10 = 45, ce qui satisfait l'équation.

Ainsi, la solution est x = 5, y = 10.

Par conséquent, le nombre de cailloux dans chaque boîte est le suivant :

- Première boîte : x = 5 cailloux
- Deuxième boîte : x + y = 5 + 10 = 15 cailloux
- Troisième boîte : 2x = 2 * 5 = 10 cailloux
- Quatrième boîte : 3x = 3 * 5 = 15 cailloux

Encadré 3 : Solution de ChatGPT (version 3.5, janvier 2024)

Lors de la troisième tentative (encadrés 4a et 4b) quatre mois plus tard et une nouvelle version de ChatGPT le compte est bon. A noter que la deuxième variable (nombre de cailloux de la deuxième boîte est notée textuellement « nombre de cailloux dans la deuxième boîte » avant d'être renommée *y*.

Lors de la résolution par tentatives, la première valeur est écartée après contrôle d'une des conditions posée dans l'énoncé. Le deuxième essai est le bon. Puis le système vérifie encore que le

total correspond. Il n'y a pas d'indication de pourquoi commencer par $x=5$.

De la première tentative à la troisième, que s'est-il passé ? Enrichissement du modèle général ? Enrichissement des capacités intégrées en *raisonnement* mathématique[3] ? Entrainement sur des problèmes isomorphes à celui des cailloux ? Nous privilégions pour le moment la troisième l'hypothèse.

ChatGPT 4o

Pour trouver combien de cailloux il y a dans chaque boîte, commençons par définir le nombre de cailloux dans la première boîte comme x .

Ensuite, nous utiliserons les informations fournies pour déterminer le nombre de cailloux dans les autres boîtes :

- Dans la première boîte, il y a x cailloux.

- Dans la deuxième boîte, il y a plus de cailloux que dans la première boîte, mais moins que dans la troisième. Nous n'avons pas encore assez d'informations pour déterminer exactement combien il y en a, alors nous reviendrons à cette boîte plus tard.

- Dans la troisième boîte, il y a le double du nombre de cailloux de la première boîte, donc il y a 2x cailloux.

- Dans la quatrième boîte, il y a le triple du nombre de cailloux de la première boîte, donc il y a 3x cailloux.

Sachant que Jacques a ramassé un total de 45 cailloux, nous pouvons écrire l'équation suivante en utilisant les informations pour la première, troisième, et quatrième boîte :

x + \text{(nombre de cailloux dans la deuxième boîte)} + 2x + 3x = 45

Simplifions cette équation : x + \text{(nombre de cailloux dans la deuxième boîte)} + 5x = 45

6x + \text{(nombre de cailloux dans la deuxième boîte)} = 45

Pour que le nombre de cailloux dans la deuxième boîte soit plus que dans la première boîte mais moins que dans la troisième, il doit se situer entre x et 2x. Soit y le nombre de cailloux dans la deuxième boîte, alors : x < y < 2x

Réécrivons l'équation pour y : 6x + y = 45 Puisque x < y < 2x , et en utilisant le fait que y doit être un entier, nous devons trouver une valeur pour x qui respecte cette inégalité et qui permet y d'être aussi un entier. Essayer des valeurs pour x :

1. Si x = 5 , alors :

6(5) + y = 45

30 + y = 45

y = 15

Vérifions si y = 15 respecte l'inégalité x < y < 2x :

5 < 15 < 10 (Faux)

Encadré 4a : Solution de ChatGPT, essai avec x=5 (version 4o, mai 2024)

3 A ce propos voir la réponse de ChatGPT à l'invite : « Quel module ou plugin utilises-tu pour les MWP » dans la chronique dédiée au problème des chameaux et des dromadaires déjà citée.

2. Si $x = 6$, alors :

$6(6) + y = 45$

$36 + y = 45$

$y = 9$

Vérifions si $y = 9$ respecte l'inégalité $x < y < 2x$:

$6 < 9 < 12$ (Vrai)

Donc, $x = 6$ fonctionne.

Nous avons trouvé que :

- La première boîte contient $x = 6$ cailloux.
- La deuxième boîte contient $y = 9$ cailloux.
- La troisième boîte contient $2x = 12$ cailloux.
- La quatrième boîte contient $3x = 18$ cailloux.

En vérifiant, nous avons : $6 + 9 + 12 + 18 = 45$

Ce qui est correct. Par conséquent, les nombres de cailloux dans chaque boîte sont :

- Première boîte : 6 cailloux
- Deuxième boîte : 9 cailloux
- Troisième boîte : 12 cailloux
- Quatrième boîte : 18 cailloux

Encadré 4b : Suite de la solution de ChatGPT (version 4o, mai 2024)

Gemini

Gemini est utilisé à deux reprises.

Dans le premier cas (encadré 5) la démarche est extrêmement bien structurée. La mathématisation est soignée de même que la formulation de la solution. La relation simplifiée est vite trouvée. Lors de l'étape 6, il (ou elle) prend pour y (l'accroissement) un multiple de x. Les tentatives portent sur la valeur de y. Il (ou elle) s'en tient au deux possibilités $x=y$ et $y=2x$ sans considérer les valeurs intermédiaires. Il n'y a pas d'indication pourquoi commencer et s'en tenir à $x=5$. Iel obtient la même solution que chatGPT (encadré 5).

Dans le deuxième essai, y est le nombre de cailloux de la deuxième boîte (et non l'accroissement). Le système ayant attribué 45 cailloux à la première boîte déraille dès la troisième étape et aboutit à une solution et une remarque hors de propos quoique bien formulée. A noter que les quantités de cailloux sont tous des multiples de 15,

Ces deux essais exécutés l'un quasiment à la suite de l'autre mettent en évidence l'aspect génératif (i.e un certain aléa dans la production du texte). La nature de l'ensemble et de la méthode d'entrainement reste un mystère à éclaircir.

Répartition des cailloux de Jacques

Étape 1 : Déterminons le nombre de cailloux dans la première boîte: Soit x le nombre de cailloux dans la première boîte.

Étape 2 : Déterminons le nombre de cailloux dans la deuxième boîte:

La deuxième boîte contient plus de cailloux que la première, mais moins que la troisième. Donc, le nombre de cailloux dans la deuxième boîte est compris entre x et $2x$.

Étape 3 : Déterminons le nombre de cailloux dans la troisième boîte:

La troisième boîte contient le double du nombre de cailloux contenus dans la première. Donc, le nombre de cailloux dans la troisième boîte est $2x$.

Étape 4 : Déterminons le nombre de cailloux dans la quatrième boîte:

La quatrième boîte contient le triple du nombre de cailloux contenus dans la première. Donc, le nombre de cailloux dans la quatrième boîte est $3x$.

Étape 5 : Écrivons l'équation et trouvons la valeur de x:

La somme des cailloux dans les quatre boîtes est de 45. Donc, $x + (x + y) + 2x + 3x = 45$. En regroupant les termes semblables, on obtient $7x + y = 45$.

Étape 6 : Déterminons la valeur de y: On sait que y est un nombre compris entre x et $2x$. En testant les valeurs possibles de y, on trouve que :

- Si $y = x$, alors $7x + y = 45$ devient $8x = 45$, ce qui donne $x = 5.625$. Ce n'est pas possible car le nombre de cailloux doit être un entier.
- Si $y = 2x$, alors $7x + y = 45$ devient $9x = 45$, ce qui donne $x = 5$.

Étape 7 : Calculons le nombre de cailloux dans chaque boîte:

- Boîte 1 : $x = 5$ cailloux.
- Boîte 2 : $x + y = 5 + 2(5) = 15$ cailloux.
- Boîte 3 : $2x = 2(5) = 10$ cailloux.
- Boîte 4 : $3x = 3(5) = 15$ cailloux.

Vérification:

La somme des cailloux dans les quatre boîtes est $5 + 15 + 10 + 15 = 45$.

Conclusion:

- Boîte 1 : 5 cailloux
- Boîte 2 : 15 cailloux
- Boîte 3 : 10 cailloux
- Boîte 4 : 15 cailloux

Encadré 5 : Première solution de Gemini (février 2024)

Déterminons le nombre de cailloux dans chaque boîte :

Étape 1 : Définir les variables

- `x` : nombre de cailloux dans la première boîte
- `y` : nombre de cailloux dans la deuxième boîte
- `z` : nombre de cailloux dans la troisième boîte

Étape 2 : Écrire les équations

D'après les informations données, nous pouvons écrire les équations suivantes :

1. **Première boîte:** `x` cailloux
2. **Deuxième boîte:** `y` cailloux, avec `x < y < z`
3. **Troisième boîte:** `z = 2x` cailloux
4. **Quatrième boîte:** `3x` cailloux

Étape 3 : Résoudre le système d'équations

3.1. Déterminer le nombre de cailloux dans la troisième boîte:

En utilisant l'équation `z = 2x`, on peut directement déduire que `z = 2 * 45 = 90` cailloux.

3.2. Déterminer le nombre de cailloux dans la quatrième boîte:

En utilisant l'équation `3x`, on peut déduire que la quatrième boîte contient `3 * 45 = 135` cailloux.

3.3. Déterminer le nombre de cailloux dans la deuxième boîte:

- On sait que `x < y < z`.
- On a `x = 45` et `z = 90`.
- Par conséquent, `45 < y < 90`.

Deux cas possibles :

Cas 1 : `y = 60`

Si `y = 60`, on a :

- `x = 45`
- `y = 60`
- `z = 2x = 90`
- `3x = 135`

Cas 2 : `y = 75`

Si `y = 75`, on a :

- `x = 45`
- `y = 75`
- `z = 2x = 90`
- `3x = 135`

Dans les deux cas, les valeurs correspondent aux informations données.

Encadré 6a : Deuxième solution de Gemini (février 2024)

Étape 4 : Conclusion

Deux solutions possibles existent :

Solution 1:

- Boîte 1 : 45 cailloux
- Boîte 2 : 60 cailloux
- Boîte 3 : 90 cailloux
- Boîte 4 : 135 cailloux

Solution 2:

- Boîte 1 : 45 cailloux
- Boîte 2 : 75 cailloux
- Boîte 3 : 90 cailloux
- Boîte 4 : 135 cailloux

Les deux solutions sont valides et correspondent aux informations données.

Remarque :

- Il est important de vérifier que les solutions trouvées correspondent aux informations données dans l'énoncé.
- Dans cet exemple, deux solutions existent car il n'y a pas d'informations supplémentaires permettant de déterminer le nombre exact de cailloux dans la deuxième boîte.

Encadré 6b : Deuxième solution de Gemini (fin)

Mistral

Deux essais sont effectués (encadrés 7a, 7b et 8). La première réponse est parfaite ! Toutefois, aucun argument n'est fournit pour éliminer les valeurs inférieures à 5 et supérieures à 7 pour x.

La solution du deuxième essai est par contre moins judicieuse. Le choix de la valeur de x est mystérieux. Suivi d'un raisonnement correct, il propose une vérification dont l'acceptation repose sur une faute de calcul. Le système ne semble pas en mesure de traiter des opérations arithmétiques autrement qu'en mode génératif (c'est-à-dire dans procéder véritablement au calcul).

Yaiho

Le 31 décembre 2024, presque une année après les essais précédents, le problème est proposé à Yiaho. La méthode tranche avec ce qui avait été observé jusque là. En effet, c'est la seule solution parmi celles proposées où le nombre de cailloux dans la première boîte est obtenu par déduction à partir d'une double inégalité (sans passer par des tentatives) (encadrés 9a et 9b).

Une nouvelle session introduit une petite perturbation : « Jacques a ramassé 51 cailloux colorés et non 45 comme signalé par erreur ». La méthode est la même que précédemment et aboutit à la solution exacte (7, 9, 14, 21), sans aucune référence à « l'erreur » signalée dans la donnée.

On sait que Jacques a ramassé 45 cailloux colorés et qu'il veut les répartir en quatre boîtes. On va utiliser les informations données pour trouver le nombre de cailloux dans chaque boîte.

Encadré 7a : Résumé de la consigne par Mistral

Soit x le nombre de cailloux dans la première boîte. Dans la deuxième boîte, il y a plus de cailloux que dans la première, mais moins que dans la troisième. Donc, on peut appeler y le nombre de cailloux dans la deuxième boîte, avec $x < y < 2x$ (puisque la troisième boîte contient le double du nombre de cailloux de la première).

Dans la troisième boîte, il y a le double du nombre de cailloux contenus dans la première, soit 2x.

Dans la quatrième boîte, il y a le triple du nombre de cailloux contenus dans la première, soit 3x.

Maintenant, on sait que la somme des cailloux dans les quatre boîtes est égale à 45 : $x + y + 2x + 3x = 45$

On peut simplifier l'équation en regroupant les termes avec x : $6x + y = 45$

On sait aussi que $x < y < 2x$, donc on peut tester les valeurs possibles de x pour trouver une solution qui satisfait ces conditions.

Si $x = 5$, alors y doit être compris entre 5 et 10 (exclus), mais il n'y a pas de nombre entier dans cet intervalle qui permet de satisfaire l'équation $6x + y = 45$. Si $x = 6$, alors y doit être compris entre 6 et 12 (exclus). Si on teste $y = 9$, on obtient : $6x + y = 6 * 6 + 9 = 45$

Cette solution satisfait toutes les conditions, donc on a trouvé le nombre de cailloux dans chaque boîte :

Première boîte : $x = 6$ cailloux
Deuxième boîte : $y = 9$ cailloux
Troisième boîte : $2x = 12$ cailloux
Quatrième boîte : $3x = 18$ cailloux

Vérification : $6 + 9 + 12 + 18 = 45$ cailloux.

Encadré 7 : Première solution de Mistral (février 2024)

D'après les informations données, nous pouvons établir les équations suivantes :

1. La première boîte contient x cailloux.
2. La deuxième boîte contient plus de cailloux que la première, mais moins que la troisième. Donc, elle contient k cailloux, où $x < k < 2x$.
3. La troisième boîte contient 2x cailloux (le double du nombre de cailloux contenus dans la première boîte).
4. La quatrième boîte contient 3x cailloux (le triple du nombre de cailloux contenus dans la première boîte).

La somme des cailloux dans les quatre boîtes est égale à 45. Donc, nous avons l'équation suivante : $x + k + 2x + 3x = 45$

Comme $x < k < 2x$, le seul nombre entier possible pour x qui satisfait cette condition est $x = 7$.

Si $x = 7$, alors k peut être 8, 9, 10, 11, 12 ou 13. Cependant, si nous essayons ces valeurs, nous trouvons que seul $k = 11$ convient, car $7 + 11 + 2x7 + 3x7 = 45$.

Donc, la première boîte contient 7 cailloux, la deuxième boîte contient 11 cailloux, la troisième boîte contient 14 cailloux (2x7), et la quatrième boîte contient 21 cailloux (3x7).

Encadré 8 : Deuxième solution de Mistral (février 2024)

Pour résoudre ce problème, définissons le nombre de cailloux dans chaque boîte :

- Soit (x) le nombre de cailloux dans la première boîte.
- Dans la deuxième boîte, Jacques a mis (y) cailloux.
- Dans la troisième boîte, il y a (z) cailloux.
- Dans la quatrième boîte, il y a (w) cailloux.

D'après les informations données :

1. (y > x) (la deuxième boîte contient plus de cailloux que la première)
2. (y < z) (la deuxième boîte contient moins de cailloux que la troisième)
3. (z = 2x) (la troisième boîte contient le double de cailloux de la première)
4. (w = 3x) (la quatrième boîte contient le triple de cailloux de la première)
5. La somme totale des cailloux est de 45 : [x + y + z + w = 45]

Maintenant, remplaçons (z) et (w) par leurs expressions en fonction de (x) :

[x + y + 2x + 3x = 45]

[6x + y = 45]

Cela nous donne l'équation suivante : [y = 45 - 6x]

Maintenant, en utilisant les conditions sur (y), nous savons que (y > x) et (y < z).

Étape 1 : Résoudre (y > x)

[45 - 6x > x] [45 > 7x] [x < \frac{45}{7} \approx 6.43]

Donc, (x) peut être au maximum 6 (puisqu'il doit être un entier).

Étape 2 : Résoudre (y < z)

Sachant que (z = 2x), nous avons : [45 - 6x < 2x] [45 < 8x] [x > \frac{45}{8} \approx 5.625]

Donc, (x) doit être au moins 6.

Conclusion sur (x)

Ainsi, la seule valeur d'entier valide pour (x) est 6.

Étape 3 : Calculer les autres valeurs

Si (x = 6) :

- Pour (z) : [z = 2x = 2 \times 6 = 12]
- Pour (w) : [w = 3x = 3 \times 6 = 18]
- Pour (y) : [y = 45 - 6x = 45 - 36 = 9]

Encadré 9a : *Déduction* permettant à Yaiho de déterminer le nombre de cailloux dans la première boîte

Étape 4 : Vérification

1. (x = 6)
2. (y = 9)
3. (z = 12)
4. (w = 18)

Vérifions les conditions :

- (y > x) (9 > 6) : vrai
- (y < z) (9 < 12) : vrai
- Somme : (6 + 9 + 12 + 18 = 45) : vrai

Résultat final

- Boîte 1 : 6 cailloux
- Boîte 2 : 9 cailloux
- Boîte 3 : 12 cailloux
- Boîte 4 : 18 cailloux

Encadré 9b : Fin de la solution de Yiaho

Deuxième problème : Collections de cailloux (cat 5-7)

Ce problème, avec 5 boîtes au lieu de 4, est semblable au précédent. Il est tiré de la banque de problèmes du Rallye mathématique transalpin (RMT)[4]. Il est de catégorie 5-7 (élèves de 11 à 13 ans).

Jacques a ramassé 57 cailloux colorés et les a répartis dans 5 boîtes.

Dans la troisième boîte, le nombre de cailloux est le double de celui de la première boîte.

Dans la deuxième boîte, il y a plus de cailloux que dans la première et moins que dans la troisième.

Dans la cinquième boîte, le nombre de cailloux est le triple de celui des cailloux de la première boîte.

Dans la quatrième boîte, il y a plus de cailloux que dans la troisième et moins que dans la cinquième.

Combien de cailloux peut-il y avoir dans chaque boîte ?

Indiquez toutes les manières de répartir les cailloux dans les boîtes et montrez comment vous les avez trouvées.

Encadré 10 : Le problème les cailloux avec 5 boîtes

4 Problème 10 de la première épreuve du 29e RMT (2022): http://www.projet-ermitage.org/ARMT/navi_fic2.php?code=op147-fr&flag=1&langue=fr&enonce=29rmti_fr-10&w=0

Esquisse d'une solution en IN

Le nombre de cailloux dans les boîtes : x, y, 2x, z, 3x

Avec les contraintes : $x < y < 2x < z < 3x$ et $6x + y + z = 57$.

Donc $6x + x + 2x < 6x + y + z = 57 < 6x + 2x + 3x$ en utilisant $x < y < 2x$ et $2x < z < 3x$

Donc $9x < 57 < 11x$ et donc $x < 7$ et $x > 5$ et donc $x = 6$. Les contrainte deviennent :

$6 < y < 12 < z < 18$ et $y + z = 57 - 36 = 21$

$z > 12$ implique $y < 9$

On essaye pour y les valeurs 7 et 8 ; les deux fonctionnent :

Solutions (6, 7, 12, 14, 18) et (6, 8, 12, 13, 18)

En moyenne, selon la catégorie, 2/3 des classes où le problème est posé obtiennent au moins une des deux solutions.

Solutions proposées par des systèmes IA

ChatGPT

Lors d'un premier essai avec la version 3.5, les couples (y, z) examinés, contrairement à ce qu'affirme le système, ne respectent pas les contraintes (encadrés 11a et 11b).

> Pour résoudre ce problème, nous pouvons utiliser des équations basées sur les informations fournies.
>
> Soit x le nombre de cailloux dans la première boîte.
>
> Ensuite, nous pouvons utiliser les informations données pour écrire les équations suivantes :
>
> 1. Dans la troisième boîte, le nombre de cailloux est le double de celui de la première boîte : Nombre de cailloux dans la troisième boîte = 2x
> 2. Dans la cinquième boîte, le nombre de cailloux est le triple de celui de la première boîte : Nombre de cailloux dans la cinquième boîte = 3x
> 3. Dans la deuxième boîte, il y a plus de cailloux que dans la première et moins que dans la troisième : Nombre de cailloux dans la première boîte < Nombre de cailloux dans la deuxième boîte < Nombre de cailloux dans la troisième boîte
> 4. Dans la quatrième boîte, il y a plus de cailloux que dans la troisième et moins que dans la cinquième : Nombre de cailloux dans la troisième boîte < Nombre de cailloux dans la quatrième boîte < Nombre de cailloux dans la cinquième boîte
>
> Enfin, nous savons que la somme totale des cailloux est égale à 57 :
>
> x + (2x) + y + z + (3x) = 57
>
> Où y est le nombre de cailloux dans la deuxième boîte et z est le nombre de cailloux dans la quatrième boîte.
>
> En résolvant cette équation, nous pouvons trouver les valeurs possibles pour x, y et z.

Encadré 11a : Solution de ChatGPT (version 3.5, janvier 2024)

Je vais calculer les solutions possibles.

Pour résoudre l'équation, nous avons :

$x + 2x + y + z + 3x = 57$

$6x + y + z = 57$

Maintenant, nous devons examiner les contraintes données :

1. $x < y < 2x$ (pour la deuxième boîte)
2. $2x < z < 3x$ (pour la quatrième boîte)

Nous devons également tenir compte du fait que x, y et z doivent être des nombres entiers, car nous parlons de cailloux.

Nous pouvons procéder par essais et erreurs pour trouver toutes les solutions possibles.

Commençons par examiner les valeurs possibles pour x. Puisque la somme totale est 57, x doit être inférieur à 10 pour que 3x (le nombre maximum de cailloux dans une boîte) soit inférieur à 30.

Essayons avec différentes valeurs de x :

1. Si x = 1 : 6 * 1 + y + z = 57 y + z = 51 Les paires possibles (y, z) pour satisfaire à la contrainte sont (13, 38), (14, 37), (15, 36), (16, 35), (17, 34), (18, 33), (19, 32), (20, 31), (21, 30). Cependant, seulement (14, 37) et (17, 34) respectent les autres contraintes.
2. Si x = 2 : 6 * 2 + y + z = 57 y + z = 45 Les paires possibles (y, z) pour satisfaire à la contrainte sont (12, 33), (13, 32), (14, 31), (15, 30). Cependant, seulement (13, 32) respecte les autres contraintes.
3. Si x = 3 : 6 * 3 + y + z = 57 y + z = 33 Les paires possibles (y, z) pour satisfaire à la contrainte sont (11, 22), (12, 21). Cependant, aucune de ces paires ne respecte les autres contraintes.

Donc, les solutions possibles sont :

- x = 1, y = 14, z = 37
- x = 1, y = 17, z = 34
- x = 2, y = 13, z = 32

Ce sont toutes les manières de répartir les cailloux dans les boîtes qui respectent les conditions données.

Encadré 11b : Solution de ChatGPT (suite) (version 3.5, janvier 2024)

Quelques mois plus tard (mai 2024) la version 4o venait à bout du problème (encadrés 12a, 12b, 12c). Diable ! cette fois-ci les deux bonnes solutions sont trouvées (et la troisième n'est pas extravagante). Il semble difficilement possible qu'un « simple » processus génératif « aléatoire » aboutisse à cette solution. Les raisons du progrès resteraient à trouver : entrainement (non supervisé, supervisé, assisté?) à partir de la version précédente ? découverte de la solution dans la banque de l'ARMT ? Modèle général introduit ?

Pour résoudre ce problème, définissons \(x \) comme le nombre de cailloux dans la première boîte.

Ensuite, nous pouvons exprimer le nombre de cailloux dans les autres boîtes en fonction de \(x \) :

- Première boîte : \(x \) cailloux.

- Deuxième boîte : \(y \) cailloux, où \(x < y < 2x \).

- Troisième boîte : \(2x \) cailloux.

- Quatrième boîte : \(z \) cailloux, où \(2x < z < 3x \).

- Cinquième boîte : \(3x \) cailloux.

Sachant que le total des cailloux est 57, nous avons l'équation suivante :

\[x + y + 2x + z + 3x = 57 \]

Simplifions cette équation :

\[6x + y + z = 57 \]

Avec les contraintes suivantes pour \(y \) et \(z \) :

- \(x < y < 2x \)

- \(2x < z < 3x \)

Nous allons essayer des valeurs pour \(x \) et vérifier les conditions pour \(y \) et \(z \).

1. Si \(x = 5 \), alors :

\[6(5) + y + z = 57 \]

\[30 + y + z = 57 \]

\[y + z = 27 \]

Les contraintes deviennent :

\[5 < y < 10 \]

\[10 < z < 15 \]

Encadré 12a : Début de la solution de ChatGPT (version 4o, mai 2024)

Essayons les combinaisons possibles :

- Si $y = 6$, alors $z = 21$ (ne respecte pas $10 < z < 15$).
- Si $y = 7$, alors $z = 20$ (ne respecte pas $10 < z < 15$).
- Si $y = 8$, alors $z = 19$ (ne respecte pas $10 < z < 15$).
- Si $y = 9$, alors $z = 18$ (ne respecte pas $10 < z < 15$).

Donc $x = 5$ ne fonctionne pas.

2. Si $x = 6$, alors :

$$6(6) + y + z = 57$$

$$36 + y + z = 57$$

$$y + z = 21$$

Les contraintes deviennent :

$$6 < y < 12$$

$$12 < z < 18$$

Essayons les combinaisons possibles :

- Si $y = 7$, alors $z = 14$ (fonctionne).
- Si $y = 8$, alors $z = 13$ (fonctionne).
- Si $y = 9$, alors $z = 12$ (fonctionne).
- Si $y = 10$, alors $z = 11$ (ne respecte pas $12 < z < 18$).
- Si $y = 11$, alors $z = 10$ (ne respecte pas $12 < z < 18$).

Donc, pour $x = 6$, nous avons les combinaisons suivantes :

- $y = 7$, $z = 14$
- $y = 8$, $z = 13$
- $y = 9$, $z = 12$

3. Si $x = 7$, alors :

$$6(7) + y + z = 57$$

$$42 + y + z = 57$$

$$y + z = 15$$

Les contraintes deviennent :

$$7 < y < 14$$

$$14 < z < 21$$

Encadré 12b : Solution de ChatGPT (suite) (version 4o, mai 2024)

Essayons les combinaisons possibles :

- Si $y = 8$, alors $z = 7$ (ne respecte pas $14 < z < 21$).
- Si $y = 9$, alors $z = 6$ (ne respecte pas $14 < z < 21$).
- Si $y = 10$, alors $z = 5$ (ne respecte pas $14 < z < 21$).
- Si $y = 11$, alors $z = 4$ (ne respecte pas $14 < z < 21$).
- Si $y = 12$, alors $z = 3$ (ne respecte pas $14 < z < 21$).
- Si $y = 13$, alors $z = 2$ (ne respecte pas $14 < z < 21$).

Donc $x = 7$ ne fonctionne pas.

4. Si $x = 8$, alors :

$$6(8) + y + z = 57$$

$$48 + y + z = 57$$

$$y + z = 9$$

Les contraintes deviennent :

$$8 < y < 16$$

$$16 < z < 24$$

Essayons les combinaisons possibles :

- Si $y = 9$, alors $z = 0$ (ne respecte pas $16 < z < 24$).

Donc $x = 8$ ne fonctionne pas.

En conclusion, les seules solutions sont pour $x = 6$:

- $x = 6$, $y = 7$, $z = 14$
- $x = 6$, $y = 8$, $z = 13$
- $x = 6$, $y = 9$, $z = 12$

Les répartitions possibles des cailloux dans les boîtes sont donc :

1. 6, 7, 12, 14, 18
2. 6, 8, 12, 13, 18
3. 6, 9, 12, 12, 18

Encadré 12c : Solution de ChatGPT (fin) (version 4o, mai 2024)

Gemini

Le problème a été présenté une fois à Gemini. Contrairement à ChatGPT une inconnue est attribuée à chaque boîte avant d'en réduire le nombre.

Répartition des cailloux de Jacques

Étape 1 : Définir les variables

- `x` : nombre de cailloux dans la première boîte
- `y` : nombre de cailloux dans la deuxième boîte
- `z` : nombre de cailloux dans la troisième boîte
- `t` : nombre de cailloux dans la quatrième boîte
- `u` : nombre de cailloux dans la cinquième boîte

Étape 2 : Traduire les informations en équations

1. **Troisième boîte : double de la première**
    - `z = 2x`
2. **Deuxième boîte : entre la première et la troisième**
    - `x < y < z`
3. **Cinquième boîte : triple de la première**
    - `u = 3x`
4. **Quatrième boîte : entre la troisième et la cinquième**
    - `z < t < u`
5. **Nombre total de cailloux**
    - `x + y + z + t + u = 57`

Étape 3 : Résoudre le système d'équations

1. **Première équation:** `z = 2x`
2. **Cinquième équation:** `u = 3x`
3. **Substitution dans l'équation du nombre total:** `x + y + 2x + t + 3x = 57`
4. **Simplification:** `6x + y + t = 57`

Encadré 13a : Solution Gemini (mai 2024)

Les quatre premières étapes sont parfaitement menées de même que la présentation des solutions erronées (encadrés 13a et 13b). Un maximum pour la première boîte est déduit rapidement, mais le plan exposé concernant les vérification n'est pas suivi d'effets. La résolution est suivie de remarques et de méthodes alternatives peu informatives. C'est un « bouchon » qu'utilisent fréquemment ces systèmes génératifs[5].

5 Correspond au « parlez-moi de vos parents » d'Eliza !

Étape 4 : Explorer les solutions possibles

1. **Valeurs possibles pour x:**
    - `x` doit être un entier positif car il représente un nombre de cailloux.
    - La valeur maximale de `x` est 9 car `6x + y + t` doit être inférieur ou égal à 57.
    - On peut donc tester les valeurs `x = 1, 2, 3, ..., 9`.
2. **Pour chaque valeur de x:**
    - Calculer `z = 2x` et `u = 3x`.
    - Trouver les valeurs possibles pour `y` et `t` en tenant compte des inégalités `x < y < z` et `z < t < u`.
    - Vérifier si `x + y + z + t + u = 57`.

Étape 5 : Solutions possibles

Voici les solutions possibles pour la répartition des cailloux :

Boîte	Solution 1	Solution 2	Solution 3	Solution 4
1	1	2	3	4
2	2	3	4	5
3	2	4	6	8
4	3	5	7	9
5	3	6	9	12

Étape 6 : Conclusion

Il y a **4 solutions possibles** pour la répartition des cailloux dans les boîtes.

Remarques

- Il est important de tester toutes les valeurs possibles de `x` pour trouver toutes les solutions.
- On peut également utiliser des méthodes mathématiques plus avancées pour résoudre ce type de problème.

Méthodes alternatives

- **Programmation informatique:** On peut écrire un programme informatique pour tester toutes les combinaisons possibles de valeurs pour `x`, `y`, `z`, `t` et `u`.
- **Tableaux et graphiques:** On peut utiliser des tableaux et des graphiques pour visualiser les différentes solutions possibles.

Encadré 13b : Solution Gemini (fin)

Mistral

A noter tout d'abord que les « inconnues » sont désignées en toute lettre. La première étape se focalise de façon mystérieuse sur un argument de divisibilité de 57. Iel en *déduit* que le nombre de cailloux dans le première boîte peut être 1 ou 3 (encadré 14a). Les tentatives subséquentes ne permettent pas de respecter la contrainte sur la somme (encadré 14b). Ce qui est correctement déduit. La conclusion logique qui en découle est l'absence de solution.

A noter que le système ne remet pas en cause son analyse : « Il est possible qu'il y ait une erreur dans les informations fournies ou que le problème soit impossible à résoudre avec les contraintes

données. »

Pour résoudre ce problème, nous allons procéder étape par étape en utilisant les informations données.

1. Jacques a ramassé 57 cailloux colorés et les a répartis dans 5 boîtes.
2. Dans la troisième boîte, le nombre de cailloux est le double de celui de la première boîte. (Troisième = 2 * Première)
3. Dans la deuxième boîte, il y a plus de cailloux que dans la première et moins que dans la troisième. (Première < Seconde < Troisième)
4. Dans la cinquième boîte, le nombre de cailloux est le triple de celui des cailloux de la première boîte. (Cinquième = 3 * Première)
5. Dans la quatrième boîte, il y a plus de cailloux que dans la troisième et moins que dans la cinquième. (Troisième < Quatrième < Cinquième)

Commençons par trouver le nombre de cailloux dans la première boîte. Étant donné que le nombre de cailloux dans la troisième boîte est le double de celui de la première boîte et que le nombre de cailloux dans la cinquième boîte est le triple de celui de la première boîte, nous pouvons déduire que le nombre de cailloux dans la première boîte doit être un diviseur commun de 57 qui permet d'avoir des nombres entiers de cailloux dans les boîtes 3 et 5.

Les diviseurs de 57 sont 1, 3, 19, et 57. Cependant, si nous choisissons 19 ou 57, nous aurons des nombres de cailloux supérieurs à 57 pour les boîtes 3 et 5, ce qui n'est pas possible. Donc, nous pouvons seulement choisir entre 1 et 3 comme nombre de cailloux dans la première boîte.

Encadré 14a : Solution proposée par Mistral au 2e problème des cailloux

Essayons avec 1 caillou dans la première boîte :

- Première boîte : 1 caillou
- Troisième boîte : 2 * 1 = 2 cailloux
- Cinquième boîte : 3 * 1 = 3 cailloux

Maintenant, nous devons trouver le nombre de cailloux dans la deuxième boîte. Comme il y a plus de cailloux que dans la première boîte et moins que dans la troisième, la deuxième boîte doit contenir 2 cailloux.

- Seconde boîte : 2 cailloux

Ensuite, nous devons trouver le nombre de cailloux dans la quatrième boîte. Comme il y a plus de cailloux que dans la troisième boîte et moins que dans la cinquième, la quatrième boîte doit contenir 3 cailloux.

- Quatrième boîte : 3 cailloux

Cependant, en additionnant tous ces cailloux, nous obtenons un total de 1 + 2 + 2 + 3 + 3 = 11 cailloux, ce qui est inférieur à 57. Par conséquent, la répartition avec 1 caillou dans la première boîte ne fonctionne pas.

Essayons maintenant avec 3 cailloux dans la première boîte :

- Première boîte : 3 cailloux
- Troisième boîte : 2 * 3 = 6 cailloux
- Cinquième boîte : 3 * 3 = 9 cailloux

Pour la deuxième boîte, il y a plus de cailloux que dans la première et moins que dans la troisième, donc la deuxième boîte doit contenir 4, 5, ou 6 cailloux. Cependant, si nous choisissons 6 cailloux, il n'y aura plus assez de cailloux pour les boîtes 4 et 5. Donc, nous devons choisir entre 4 et 5 cailloux pour la deuxième boîte.

- Seconde boîte : 4 ou 5 cailloux

Pour la quatrième boîte, il y a plus de cailloux que dans la troisième et moins que dans la cinquième, donc la quatrième boîte doit contenir 7 ou 8 cailloux.

- Quatrième boîte : 7 ou 8 cailloux

Si nous choisissons 4 cailloux pour la deuxième boîte et 8 cailloux pour la quatrième boîte, nous obtenons un total de 3 + 4 + 6 + 8 + 9 = 30 cailloux, ce qui est inférieur à 57. Par conséquent, cette répartition ne fonctionne pas non plus.

Si nous choisissons 5 cailloux pour la deuxième boîte et 7 cailloux pour la quatrième boîte, nous obtenons un total de 3 + 5 + 6 + 7 + 9 = 30 cailloux, ce qui est également inférieur à 57. Par conséquent, cette répartition ne fonctionne pas non plus.

Après avoir essayé toutes les possibilités, il semble qu'il n'y ait aucune manière de répartir les 57 cailloux dans les boîtes en respectant les contraintes données. Il est possible qu'il y ait une erreur dans les informations fournies ou que le problème soit impossible à résoudre avec les contraintes données.

Encadré 14b : Fin de la solution proposée par Mistral

Yiaho

Le système ne suit pas complètement la voie royale qu'il a tracée pour la résolution du problème avec quatre boîtes. Il n'utilise que partiellement les inégalités (un usage optimal demande deux étapes) pour donner 1 et 9 comme limites pour le nombre de cailloux dans la première boîte (encadré 15a). Ces valeurs sont systématiquement passées en revue (encadré 15b). Les valeurs de 3 à 8 sont jugées possibles. Sans autre déduction la valeur juste est proposée mais le nombre de cailloux dans la quatrième boîte ne satisfait pas l'inégalité souhaitée contrairement à ce qu'affirme le commentaire : « C'est une des nombreuses manières de répartir les cailloux, mais c'est une solution valide qui respecte toutes les conditions. » (encadré 15c).

A noter l'utilisation d'indices dans la notation des inconnues.

Pour résoudre ce problème, définissons les variables suivantes :

- Soit (x_1) le nombre de cailloux dans la première boîte.
- Soit (x_2) le nombre de cailloux dans la deuxième boîte.
- Soit (x_3) le nombre de cailloux dans la troisième boîte.
- Soit (x_4) le nombre de cailloux dans la quatrième boîte.
- Soit (x_5) le nombre de cailloux dans la cinquième boîte.

Nous avons alors les équations et inégalités suivantes basées sur les informations fournies :

1. (x_1 + x_2 + x_3 + x_4 + x_5 = 57) (total de cailloux)
2. (x_3 = 2x_1) (la troisième boîte a le double de la première)
3. (x_2 > x_1) et (x_2 < x_3) (la deuxième boîte a plus de cailloux que la première et moins que la troisième)
4. (x_5 = 3x_1) (la cinquième boîte a le triple de la première)
5. (x_4 > x_3) et (x_4 < x_5) (la quatrième boîte a plus de cailloux que la troisième et moins que la cinquième)

Étape 1 : Remplacer (x_3) et (x_5)

Substituons (x_3) et (x_5) dans la première équation :

[x_1 + x_2 + 2x_1 + x_4 + 3x_1 = 57]

Cela se simplifie à :

[6x_1 + x_2 + x_4 = 57]

Étape 2 : Isoler (x_2)

[x_2 + x_4 = 57 - 6x_1]

Étape 3 : Déterminer les limites pour (x_1)

Pour que (x_2) et (x_4) soient des nombres positifs, nous devons avoir :

[57 - 6x_1 > 0 \implies x_1 < 9.5]

Comme (x_1) doit être un entier, (x_1) peut prendre les valeurs de 1 à 9.

Encadré 15a : Trois premières étapes de la solution proposée par Yiaho (décembre 2024)

Étape 4 : Tester les valeurs possibles pour (x_1)

1. Pour (x_1 = 9) :

[x_3 = 18, \quad x_5 = 27] [x_2 + x_4 = 57 - 54 = 3] *Impossible* car (x_2 > 9) et (x_4 < 27).

2. Pour (x_1 = 8) :

[x_3 = 16, \quad x_5 = 24] [x_2 + x_4 = 57 - 48 = 9] *Possible* :

- Si (x_2 = 9, x_4 = 0) (mais (x_4 > x_3) donc impossible).

3. Pour (x_1 = 7) :

[x_3 = 14, \quad x_5 = 21] [x_2 + x_4 = 57 - 42 = 15] *Possible* :

- Si (x_2 = 8, x_4 = 7) (mais impossible car (x_4 < x_3)).

4. Pour (x_1 = 6) :

[x_3 = 12, \quad x_5 = 18] [x_2 + x_4 = 57 - 36 = 21] *Possible* :

- Si (x_2 = 11, x_4 = 10).

5. Pour (x_1 = 5) :

[x_3 = 10, \quad x_5 = 15] [x_2 + x_4 = 57 - 30 = 27] *Possible* :

- Si (x_2 = 11, x_4 = 16) ou d'autres combinaisons.

6. Pour (x_1 = 4) :

[x_3 = 8, \quad x_5 = 12] [x_2 + x_4 = 57 - 24 = 33] *Possible* :

- (x_2 = 9, x_4 = 24).

7. Pour (x_1 = 3) :

[x_3 = 6, \quad x_5 = 9] [x_2 + x_4 = 57 - 18 = 39] *Possible* :

- (x_2 = 12, x_4 = 27).

8. Pour (x_1 = 2) :

[x_3 = 4, \quad x_5 = 6] [x_2 + x_4 = 57 - 12 = 45] *Impossible*.

9. Pour (x_1 = 1) :

[x_3 = 2, \quad x_5 = 3] [x_2 + x_4 = 57 - 6 = 51] *Impossible*.

Encadré 15b : Les tentatives de Yiaho

Conclusion :

Nous avons une répartition valide pour :

- (x_1 = 6)
- (x_2 = 11)
- (x_3 = 12)
- (x_4 = 8)
- (x_5 = 18)

Une solution est :

- **Boîte 1 : 6 cailloux**
- **Boîte 2 : 11 cailloux**
- **Boîte 3 : 12 cailloux**
- **Boîte 4 : 8 cailloux**
- **Boîte 5 : 18 cailloux**

C'est une des nombreuses manières de répartir les cailloux, mais c'est une solution valide qui respecte toutes les conditions.

Encadré 15c : Solution proposée par Yiaho

Discussion

Résumé des solutions

ChatGPT

ChatGPT lors de sa sortie grand public peut mathématiser le problème mais *bute* sur un système d'équation non déterminé (d'autres contraintes, inégalité, divisibilité, sont à prendre en compte) (encadré 2). Une année plus tard, il *tente* de terminer la résolution en procédant à des essais. Il se *contente* de considérer qu'une contrainte égalitaire (encadré 3). Le deuxième problème est bien mathématisé (encadrés 11a et 11b). Avec une justification incomplète, il est déduit que x doit être inférieur à 10. Mais seuls les valeurs 1, 2 et 3 sont testées par la suite. Les contraintes inégalitaires sont traitées mystérieusement et incomplètement (la valeur 10 semble perturber le système).

Quatre mois plus tard dans une nouvelle tentative avec une nouvelle version la contrainte inégalitaire est prise en compte et les tentatives aboutissent au bon résultat pour le premier problème (encadrés 4a et 4b). Le deuxième problème est également résolu mais sans économie de moyens (encadrés 12a, 12b et 12c) ; les contraintes sont largement passées en revue. Il n'y a pas d'information concernant le choix de la première valeur de x (nombre de cailloux dans la première boîte) testée (5).

Gemini

Gemini est utilisé à deux reprises pour le premier problème

Dans le premier cas (encadré 5) la démarche est extrêmement bien structurée. La mathématisation est soignée de même que la formulation de la solution. La relation simplifiée est vite trouvée. En cours de route, la deuxième inconnue considérée, y, le nombre de cailloux supplémentaire, est posée

comme un multiple de x. Les tentatives s'effectuent sur y en ne considérant que les deux possibilités $x=y$ et $y=2x$ sans considérer les valeurs intermédiaires. Il n'y a pas d'indication pourquoi commencer et s'en tenir à $x=5$. La résolution se termine en proposant la même solution, fausse, que chatGPT a obtenue à la même époque,

Lors du deuxième essai (encadrés 6a et 6b) la variable y représente le nombre de cailloux de la deuxième boîte (et non l'accroissement). Le système ayant attribué 45 cailloux à la première boîte déraille dès la troisième étape et aboutit à une solution extravagante et une remarque hors de propos quoique bien formulée.

Ces deux essais exécutés l'un quasiment à la suite de l'autre mettent en évidence l'aspect génératif. La nature de l'ensemble d'entrainement reste un mystère à éclaircir.

En janvier 2025, la même solution que lors du premier essai est proposée avec la remarque : « Il est possible qu'il existe d'autres solutions, mais cette méthode par essais-erreurs nous a permis de trouver une solution qui correspond aux conditions du problème. ».

Pour le deuxième problème (encadrés 13a et 13b) la mathématisation est correctement effectuée. Le nombre de cailloux dans la première boîte doit être inférieurs à 9. Quatre solutions sont ensuite proposées sans autre explication, solutions qui ne respectent pas les contraintes imposées, notamment celle concernant la somme. La résolution est suivie de remarques et de méthodes alternatives peu informatives (phrases passe-partout). C'est un procédé que présente souvent ces systèmes génératifs.

Mistral

Deux essais sont effectués (encadrés 7 et 8). La première réponse est parfaite ! Toutefois aucun argument n'est proposé pour limiter les tentatives Il n'y a pas d'argument formulé pour limiter les tentatives de x (nombre de cailloux dans le première boîte) aux valeurs 5 et 6.

La solution du deuxième essai est par contre moins judicieuse. Le choix de la valeur de x est mystérieux. Suivi d'un raisonnement correct, il propose une vérification dont l'acceptation repose sur une faute de calcul. Le système semble traiter des opérations arithmétiques en mode génératif sans effectuer le calcul proprement dit.

Pour le deuxième problème à noter tout d'abord que les « inconnues » sont désignées en toute lettre et simplement pour résumé les données. Par la suite toutes les contraintes sont exprimées en langage courant. La première étape se concentre de façon inexpliquée sur un argument de divisibilité de 57 dont on déduit que le nombre de cailloux dans le première boîte peut être 1 ou 3 (encadré 14a). Les tentatives subséquentes ne permettent pas de respecter la contrainte sur la somme (encadré 14b). Ce qui est correctement déduit. La conclusion logique qui en découle est l'absence de solution.

Yiaho[6]

Au moment de rédiger cette chronique, presque une année plus tard que les premiers essais, le problème est proposé à Yiaho. La méthode tranche avec ce qui avait été observé jusque là (encadrés 9a et 9b).

En effet, le nombre de cailloux dans la première boîte est obtenu par déduction à partir d'une double inégalité (sans passer par des tentatives).

Une nouvelle session introduit une petite perturbation : « Jacques a ramassé 51 cailloux colorés et non 45 comme signalé par erreur ». La méthode est la même que précédemment et aboutit à la solution (7, 9, 14, 21), sans référence à « l'erreur » signalée.

6 Il nous reste à éclaircir le statut de Yiaho et son rapport avec le modèle utilisé par ChtaGPT.

Pour le deuxième problème (encadrés 15a, 15b, 15c), le système ne suit pas la voie royale tracée pour la résolution du problème avec quatre boîtes. Tout d'abord il n'utilise que partiellement les inégalités (un usage optimal demande deux étapes) pour donner 1 et 9 comme limites pour le nombre de cailloux dans la première boîte. Les valeurs de 3 à 8 sont jugées possibles sans être éliminées explicitement, seule la valeur juste (6 cailloux dans le première boîte) est proposée. Le nombre de cailloux dans la quatrième boîte ne satisfait pas l'inégalité souhaitée contrairement à ce qu'indique le « bouchon » : « C'est une des nombreuses manières de répartir les cailloux, mais c'est une solution valide qui respecte toutes les conditions ».

Pour résumé ce résumé ChatGPT au fil des versions parvient à résoudre le premier problème, puis les deuxième. Gemini n'arrive pas à la solution pour aucun des problèmes et ne montre aucune tendance évolutive. Mistral réussit au premier problème mais pas au second. Quant à Yiaho, interrogé plus tardivement, il montre davantage de *capacités* à extraire des informations d'inégalités pour le premier problème mais qu'il ne met pas en oeuvre pour le deuxième.

Finalement

Comme dans notre travail précédent nous n'avons de plan d'expérience très structuré : « toutes choses égales par ailleurs ». Cependant nous pouvons constater une certaine amélioration des résultats en fonction du temps. Nous constatons par ailleurs :

- De nombreuses formulations elliptiques : « Si nous essayons les valeurs nous constatons que ... ».

- A de rares exceptions près, les inconnues sont désignées de façon standard (x, y, ...) contrairement à ce nous avions constaté lors de notre travail avec le problème de chameaux et des dromadaires où d'autres notations étaient utilisées, notamment les initiales c et d. Nouveauté : Yiaho utilise des variables indicées pour le deuxième problème.

- Même les solutions erronées sont parsemées de remarque pertinentes, par exemple sur les valeurs limite du nombre de cailloux dans la première boîte.

- Des plans cohérents suivis d'une mise en oeuvre qui laisse à désirer (nombre d'essais insuffisant ou choix inexpliqués).

- Des conclusions pertinentes à partir de procédés défectueux.

- Des remarques finales un peu passe-partout ou alors des affirmations péremptoires démenties par les faits. Selon le cas, ces remarques peuvent être perçues comme *arrogantes* ou alors *pathétiques*.

Globalement la réussite est moindre que pour le problème des chameaux et des dromadaires qui sans recherche exhaustive sur internet semble correspondre à un modèle de tâche plus souvent rencontré. Reste la question concernant l'utilisation de procédés génératifs généraux, de méta-modèles mathématiques ou encore de « copies » de solution. Les systèmes plutôt bavards restent assez flous à ce sujet à l'exemple de Yiaho :

> Question à YIAHO : Où as-tu trouvé la solution?
>
> YIAHO IQ 200 : La solution a été trouvée en appliquant une approche systématique et analytique pour résoudre le problème mathématique posé. Voici les étapes clé que j'ai suivies:
>
> 1. **Définition des variables** : J'ai défini des variables pour représenter le nombre de cailloux dans chaque boîte afin de pouvoir exprimer les relations entre elles mathématiquement.
> 2. **Équations basées sur les conditions** : J'ai formulé des équations basées sur les informations données dans le problème, telles que les inégalités et les relations entre le nombre de cailloux dans les boîtes.

3. **Résolution des équations** : J'ai ensuite résolu ces équations pour déterminer les valeurs possibles des cailloux dans chaque boîte, tout en respectant les restrictions imposées par les conditions du problème.
4. **Vérification des résultats** : Finalement, j'ai vérifié que les valeurs trouvées respectaient toutes les conditions énoncées dans le problème.

Cette méthode permet d'atteindre une solution précise et logique en utilisant des principes mathématiques. Si vous avez d'autres questions ou souhaitez explorer un autre sujet, n'hésitez pas à demander !

De la première tentative à la troisième, que s'est-il passé ? Enrichissement du modèle général ? Enrichissement des capacités intégrées en *raisonnement* mathématique[7] ? Entrainement sur des problèmes isomorphes à celui des cailloux ? Nous privilégions la troisième l'hypothèse.

De façon plus précise, notre hypothèse est que le système ayant repéré le type de problème (sans invite particulière ; on verra dans une autre chronique que ce repérage peut receler quelques failles) va suivre les étapes selon un modèle imposé. Chaque étape peut faire appel à certaines actions basées sur des entrainements préalables ou des annotations humaines. On peut imaginer que le modèle et les étapes subissent des améliorations incrémentales par renforcement voire ajout de nouveau « scénario ». Par exemple dans le cas des cailloux, cela peut concerner le traitement des relations inégalitaires. Impossible d'imaginer un procédé génératif pure lié à un modèle de langage alignant la suite de déductions observée dans les encadrés 12. Toutefois, le problème lui-même n'a pas été enseigné par un « expert » puisque les solutions tiennent en quelques lignes. L'utilisation d'un raisonnement sur les inégalités évitent de procéder à des tentatives. Les meilleures systèmes dans leurs meilleures tentatives atteignent donc sur ces problèmes le niveau d'élèves de 9 ans pour l'un, 11 ans pour l'autre. Le niveau « expert » caractérisé par le traitement algébrique des inégalités est observé une fois avec la version simple du problème (quatre boîtes).

Pour le moment, il nous manque quelques connaissances sur les techniques de bases. Nos analogies font surtout appel à des vieilles techniques : règles de production, théorie des scripts développée notamment par Roger Schank (1977) et bien présentée par Sabah (1988), etc[8].

S'il n'y a pas de miracle, reste néanmoins les exploits :

- Identification des inconnues (valeur et nom) et des relations. D'une façon générale le passage du langage courant à une certaine standardisation (par exemple en Latex comme l'indique Hendrycks et ali (2021) dans l'élaboration du dataset MATH) ;

- Mise en oeuvre du modèle ;

- Formulation de la solution ;

- Le fonction des deux interfaces de l'identification vers le modèle et des solutions vers leur expressions ;

- Mémorisation de la situation pour des requêtes ultérieures (comme la question posée à Yiaho ci-dessus) ;

- La quantité extravagante de données exploitées et mises en oeuvre.

Ces exploits peuvent être ternis par le décalage entre la forme « parfaite », le discours plutôt convenu et le contenu plus aléatoire. Mais tout évolue à une vitesse grand « V ». Nous allons poursuivre nos observations.

7 A ce propos voir aussi la réponse de ChatGPT à l'invite : « Quel module ou plugin utilises-tu pour les MWP » dans la chronique dédiée au problème des chameaux et des dromadaires déjà citée.

8 Dans cet ordre d'idées, les processus d'entrainement correspondent à l'alors extraction des connaissances auprès d'expert.

Bibliographie

Roger C. Schank, Robert P. Abelson (1977). *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures.* Psychology Press (Artificial Intelligence Series)

Gérard Sabah (1988). *L'intelligence artificielle et le langage*. Paris : Hermès

Dan Hendrycks, Collin Burns, Saurav Kadavath, Akul Arora, Steven Basart, Eric Tang, Dawn Song, Jacob Steinhardt (2021). *Measuring Mathematical Problem Solving With the MATH Dataset* (2103.03874v2.pdf)

Alain Favre, Luc-Olivier Pochon (2024). *Systèmes génératifs et résolution de problèmes mathématiques scolaires : une chronique. Le problème des chameaux et des dromadaires* http://www.projet-ermitage.org/thema2/Chronique-chameaux.pdf

LOP & AF, janvier 2025