

Le problème des cailloux : complément

Version 4 juin 2025

Le problème des cailloux [1] a été repris début mai 2025 et soumis à Gemini en posant quelques contraintes. Le première consiste à utiliser le format « Peano » [2] et la deuxième en imposant un travail sur les inégalités.

Contrainte du format

Voici le premier prompt : *Voici 2 exemples de la façon dont je voudrais que tu résolve un problème.* Les exemples sont proposés dans l'encadré 1.

Q : Olivia a 23 €. Elle a acheté cinq bagels à 3 € chacun. Combien lui reste-t-il ?

Solution Peano :

Soit a le montant d'argent avec lequel Olivia a commencé $[[\text{var } a]]$. On nous donne $[[\text{eq } a = 23]]$.

Soit b le nombre de bagels qu'elle a achetés $[[\text{var } b]]$. On nous donne $[[\text{eq } b = 5]]$.

Soit c le prix de chaque bagel $[[\text{var } c]]$. On nous donne $[[\text{eq } c = 3]]$.

Soit d le montant total d'argent qu'elle a dépensé en bagels $[[\text{var } d]]$. Nous avons $[[\text{eq } d = b * c]]$.

Soit e le montant d'argent qui lui reste $[[\text{var } e]]$. On a $[[\text{eq } e = a - d]]$.

La réponse est la valeur de e $[[\text{réponse } e]]$.

Q : Michael avait 58 balles de golf. Mardi, il en a perdu 23. Mercredi, il en a perdu 2 de plus. Combien de balles de golf avait-il à la fin de la journée de mercredi ?

Solution Peano :

Soit a le nombre de balles de golf avec lesquelles Michael a commencé $[[\text{var } a]]$. On nous donne $[[\text{eq } a = 58]]$.

Soit b le nombre de balles de golf qu'il a perdues mardi $[[\text{var } b]]$. On nous donne $[[\text{eq } b = 23]]$.

Soit c le nombre de balles de golf qu'il a perdues mercredi $[[\text{var } c]]$. On nous donne $[[\text{eq } c = 2]]$.

Soit d le nombre de balles de golf qui lui restait $[[\text{var } d]]$. On a $[[\text{eq } d = a - b - c]]$.

La réponse est la valeur de d $[[\text{réponse } d]]$.

Encadré 1 : Les exemples de l'invite

Le prompt est complété par l'énoncé du problèmes : *Et voici le problème à résoudre* (suit le problème avec 5 boîtes. La réponse de Gemini fait l'objet des encadrés 2 et 3.

Solution Peano :

Soit t le nombre total de cailloux ramassés par Jacques $[[\text{var } t]]$. On nous donne $[[\text{eq } t = 57]]$.

Soit n le nombre de boîtes $[[\text{var } n]]$. On nous donne $[[\text{eq } n = 5]]$.

Soit b_1 le nombre de cailloux dans la première boîte $[[\text{var } b_1]]$.

Soit b_2 le nombre de cailloux dans la deuxième boîte $[[\text{var } b_2]]$.

Soit b_3 le nombre de cailloux dans la troisième boîte $[[\text{var } b_3]]$.

Soit b_4 le nombre de cailloux dans la quatrième boîte $[[\text{var } b_4]]$.

Soit b_5 le nombre de cailloux dans la cinquième boîte $[[\text{var } b_5]]$.

Nous avons l'équation de la somme totale : $[[\text{eq } b_1 + b_2 + b_3 + b_4 + b_5 = t]]$.

Relation pour la troisième boîte : $[[\text{eq } b_3 = 2 * b_1]]$.

Relation pour la deuxième boîte : $[[\text{eq } b_1 < b_2]]$ et $[[\text{eq } b_2 < b_3]]$. Ceci peut être combiné en $[[\text{eq } b_1 < b_2 < b_3]]$.

Relation pour la cinquième boîte : $[[\text{eq } b_5 = 3 * b_1]]$.

Relation pour la quatrième boîte : $[[\text{eq } b_3 < b_4]]$ et $[[\text{eq } b_4 < b_5]]$. Ceci peut être combiné en $[[\text{eq } b_3 < b_4 < b_5]]$.

Toutes les variables b_1, b_2, b_3, b_4, b_5 doivent être des entiers positifs. $[[\text{eq } b_1 > 0]]$, $[[\text{eq } b_2 > 0]]$, $[[\text{eq } b_3 > 0]]$, $[[\text{eq } b_4 > 0]]$, $[[\text{eq } b_5 > 0]]$.

En substituant b_3 et b_5 dans l'équation de la somme totale :

$$[[\text{eq } b_1 + b_2 + (2 * b_1) + b_4 + (3 * b_1) = 57]]$$

En simplifiant : $[[\text{eq } 6 * b_1 + b_2 + b_4 = 57]]$.

Nous avons aussi les contraintes combinées :

$$[[\text{eq } b_1 < b_2 < 2 * b_1]]$$

$$[[\text{eq } 2 * b_1 < b_4 < 3 * b_1]]$$

Encadré 2 : Mise en forme des données

On s'aperçoit que le système a parfaitement repris le modèle des exemples. Il code également le nombre de boîtes qui n'est pas une variable qui va intervenir par la suite.

Dans la suite de la solution, il n'est pas précisé comment affirmation que b_1 doit être au moins 6 est établie (sans tester les valeurs inférieures). C'est une chaîne de raisonnement non explicitée selon l'expression reprise de Sebastian Raschka¹. Par contre le reste du raisonnement est parfaitement explicité. Il faut noter que la notation Peano est chaque ajoutée sans que cela participe à la résolution.

En définitive, le modèle de mise en forme proposé est parfaitement *compris* et réutilisé sans que cela déclenche un mode particulier de résolution (usage de sympy, par exemple).

1 <https://magazine.sebastianraschka.com/p/understanding-reasoning-llms>

Nous devons trouver les valeurs entières de b_1, b_2, b_4 qui satisfont $6 * b_1 + b_2 + b_4 = 57$ ainsi que les inégalités ci-dessus.

En testant les valeurs possibles pour b_1 (sachant que b_1 doit être au moins 6 pour que les intervalles $]b_1, 2*b_1[$ et $]2*b_1, 3*b_1[$ contiennent au moins un entier chacun et que la somme puisse approcher 57) :

Si $b_1 = 6$:

$$b_3 = 2 * 6 = 12$$

$$b_5 = 3 * 6 = 18$$

L'équation devient : $6 * 6 + b_2 + b_4 = 57 \Rightarrow 36 + b_2 + b_4 = 57 \Rightarrow [b_2 + b_4 = 21]$.

Les inégalités deviennent : $6 < b_2 < 12$ et $12 < b_4 < 18$.

Nous cherchons des paires (b_2, b_4) qui satisfont ces conditions et dont la somme est 21.

Si $b_2 = 7$, alors $b_4 = 21 - 7 = 14$. Conditions : $6 < 7 < 12$ (Vrai) et $12 < 14 < 18$ (Vrai).

Solution 1: $(b_1, b_2, b_3, b_4, b_5) = (6, 7, 12, 14, 18)$.

Si $b_2 = 8$, alors $b_4 = 21 - 8 = 13$. Conditions : $6 < 8 < 12$ (Vrai) et $12 < 13 < 18$ (Vrai).

Solution 2: $(b_1, b_2, b_3, b_4, b_5) = (6, 8, 12, 13, 18)$.

Si $b_2 = 9$, alors $b_4 = 21 - 9 = 12$. Condition $12 < 12 < 18$ est Fausse.

Si $b_2 = 10$, alors $b_4 = 21 - 10 = 11$. Condition $12 < 11 < 18$ est Fausse.

Si $b_2 = 11$, alors $b_4 = 21 - 11 = 10$. Condition $12 < 10 < 18$ est Fausse.

Si $b_1 = 7$:

$$6 * 7 + b_2 + b_4 = 57 \Rightarrow 42 + b_2 + b_4 = 57 \Rightarrow b_2 + b_4 = 15.$$

Inégalités : $7 < b_2 < 14$ et $14 < b_4 < 21$.

Minimum $b_2 = 8$. Minimum $b_4 = 15$. Somme minimum = $8 + 15 = 23$. Comme $23 > 15$, il n'y a pas de solution pour $b_1 = 7$.

Des valeurs plus élevées de b_1 ne donneront pas non plus de solutions.

Les solutions possibles sont les combinaisons trouvées.

Solution 1 : Première boîte=6, Deuxième boîte=7, Troisième boîte=12, Quatrième boîte=14, Cinquième boîte=18. $[[var\ s1 = [6, 7, 12, 14, 18]]]$

Solution 2 : Première boîte=6, Deuxième boîte=8, Troisième boîte=12, Quatrième boîte=13, Cinquième boîte=18. $[[var\ s2 = [6, 8, 12, 13, 18]]]$

La réponse est l'ensemble des solutions possibles pour $(b_1, b_2, b_3, b_4, b_5)$. $[[réponse\ [s1, s2]]]$

Encadré 3 : Solution au problème des cailloux

Contrainte sur la méthode.

Après l'exposé du problème, la demande précise : *Peux-tu trouver la solution en utilisant les propriétés des inégalités*

Absolument ! C'est un excellent problème pour utiliser les inégalités.

Notons B_1, B_2, B_3, B_4, B_5 le nombre de cailloux dans chaque boîte respectivement.

1. Traduction des informations en équations et inégalités :

- Total de cailloux : $B_1 + B_2 + B_3 + B_4 + B_5 = 57$ (Équation 1)
- Troisième boîte : $B_3 = 2 * B_1$ (Équation 2)
- Deuxième boîte : $B_1 < B_2 < B_3$ (Inégalité 1)
- Cinquième boîte : $B_5 = 3 * B_1$ (Équation 3)
- Quatrième boîte : $B_3 < B_4 < B_5$ (Inégalité 2)
- Les nombres de cailloux sont des entiers positifs.

2. Simplification à l'aide des équations :

Remplaçons B_3 et B_5 dans les inégalités et l'équation totale en fonction de B_1 :

- Inégalité 1 devient : $B_1 < B_2 < 2*B_1$
- Inégalité 2 devient : $2B_1 < B_4 < 3B_1$
- Équation 1 devient : $B_1 + B_2 + (2B_1) + B_4 + (3B_1) = 57$
Ce qui se simplifie en : $6*B_1 + B_2 + B_4 = 57$ (Équation 4)

3. Utilisation des propriétés des inégalités pour borner B_1 :

De l'inégalité $B_1 < B_2 < 2*B_1$, puisque B_1 et B_2 sont des entiers :

- La plus petite valeur possible pour B_2 est $B_1 + 1$.
- La plus grande valeur possible pour B_2 est $2B_1 - 1$.
(Pour que B_2 existe, il faut que $B_1 + 1 \leq 2B_1 - 1$, ce qui implique $2 \leq B_1$. Donc B_1 doit être au moins 2.)

De l'inégalité $2B_1 < B_4 < 3B_1$, puisque B_1 et B_4 sont des entiers :

- La plus petite valeur possible pour B_4 est $2*B_1 + 1$.
- La plus grande valeur possible pour B_4 est $3B_1 - 1$.
(Pour que B_4 existe, il faut que $2B_1 + 1 \leq 3B_1 - 1$, ce qui implique $2 \leq B_1$. Ceci est cohérent.)

Encadré 4 : Transcription des données

Après une remarque un peu gratuite, symbolise les contraintes et procède à quelques simplifications (Encadré 4, points 1 et 2). Il s'engage ensuite à borner inférieurement B_1 par 2 (nombre de cailloux de la première boîte) : beaucoup de manipulation pour expliquer qu'entre 1 et 2 impossible de placer un nombre entier. Ce résultat sera « confirmé » par la suite mais pas utilisé.

En imaginant B2 prendre la valeur minimale possible (B1+1) et la valeur maximale B4 (Encadré 5). La valeur de B1 est déterminée. Si la démarche ne manque pas d'*imagination*, les valeurs en présence mènent immédiatement à la bonne double inégalité.

Maintenant, utilisons ces bornes dans l'Équation 4 ($6*B1 + B2 + B4 = 57$) :

● **Borne inférieure pour la somme (pour trouver une limite supérieure pour B1)**

Si B2 prend sa valeur minimale (B1 + 1) et B4 prend sa valeur minimale ($2B1 + 1$) :

$$6B1 + (B1 + 1) + (2B1 + 1) \leq 57$$

$$9B1 + 2 \leq 57$$

$$9*B1 \leq 55$$

$$B1 \leq 55/9$$

$$B1 \leq 6.11...$$

Puisque B1 est un entier, $B1 \leq 6$.

● **Borne supérieure pour la somme (pour trouver une limite inférieure pour B1)**

Si B2 prend sa valeur maximale ($2B1 - 1$) et B4 prend sa valeur maximale ($3B1 - 1$) :

$$6B1 + (2B1 - 1) + (3B1 - 1) \geq 57$$

$$11B1 - 2 \geq 57$$

$$11*B1 \geq 59$$

$$B1 \geq 59/11$$

$$B1 \geq 5.36...$$

Puisque B1 est un entier, $B1 \geq 6$.

4. Détermination de B1 :

Nous avons $B1 \leq 6$ et $B1 \geq 6$. La seule valeur entière possible pour B1 est donc : **B1 = 6**

(Ceci respecte bien la condition $B1 \geq 2$ trouvée précédemment).

Encadré 5 : Détermination de B1

Les contraintes sont ensuite ajustées avec la valeur de B1 obtenue (Encadré 6). Les différentes valeurs possibles pour B2 sont testées. Finalement, les deux solutions possibles sont présentées suivies d'une vérification (Encadré 7). Par contre, il n'est pas clair de savoir si par sa démarche le système a exclu *consciemment* la valeur $B1=7$?

5. Calcul des autres valeurs :

$$B1 = 6$$

$$B3 = 2 * B1 = 2 * 6 = 12$$

$$B5 = 3 * B1 = 3 * 6 = 18$$

Maintenant, utilisons ces valeurs dans les inégalités pour B2 et B4 :

$$\text{Pour } B2 : B1 < B2 < B3 \Rightarrow 6 < B2 < 12$$

Les valeurs entières possibles pour B2 sont : 7, 8, 9, 10, 11.

$$\text{Pour } B4 : B3 < B4 < B5 \Rightarrow 12 < B4 < 18$$

Les valeurs entières possibles pour B4 sont : 13, 14, 15, 16, 17.

$$\text{Reprenons l'Équation 4 : } 6*B1 + B2 + B4 = 57$$

$$\text{Avec } B1 = 6 :$$

$$36 + B2 + B4 = 57$$

$$B2 + B4 = 57 - 36$$

$$B2 + B4 = 21$$

Nous devons trouver des paires (B2, B4) telles que :

$$6 < B2 < 12$$

$$12 < B4 < 18$$

$$B2 + B4 = 21$$

Testons les valeurs possibles pour B2 :

$$\text{Si } B2 = 7 : B4 = 21 - 7 = 14.$$

Vérifions les conditions : $6 < 7 < 12$ (OK) et $12 < 14 < 18$ (OK). C'est une solution.

$$\text{Si } B2 = 8 : B4 = 21 - 8 = 13.$$

Vérifions les conditions : $6 < 8 < 12$ (OK) et $12 < 13 < 18$ (OK). C'est une solution.

$$\text{Si } B2 = 9 : B4 = 21 - 9 = 12.$$

Vérifions les conditions : $6 < 9 < 12$ (OK) mais $12 < 12 < 18$ (NON, car B4 doit être strictement supérieur à 12).

$$\text{Si } B2 = 10 : B4 = 21 - 10 = 11.$$

Vérifions les conditions : $6 < 10 < 12$ (OK) mais $12 < 11 < 18$ (NON).

$$\text{Si } B2 = 11 : B4 = 21 - 11 = 10.$$

Vérifions les conditions : $6 < 11 < 12$ (OK) mais $12 < 10 < 18$ (NON).

Encadré 6 : Détermination des autres inconnues.

6. Solutions possibles :

Il y a deux combinaisons possibles pour le nombre de cailloux dans chaque boîte :

Solution 1 :

- Boîte 1 (B1) : 6 cailloux
- Boîte 2 (B2) : 7 cailloux
- Boîte 3 (B3) : 12 cailloux ($2 * 6$)
- Boîte 4 (B4) : 14 cailloux
- Boîte 5 (B5) : 18 cailloux ($3 * 6$)

Vérification du total : $6 + 7 + 12 + 14 + 18 = 57$.

Vérification des inégalités : $6 < 7 < 12$ (OK) et $12 < 14 < 18$ (OK).

Solution 2 :

- Boîte 1 (B1) : 6 cailloux
- Boîte 2 (B2) : 8 cailloux
- Boîte 3 (B3) : 12 cailloux ($2 * 6$)
- Boîte 4 (B4) : 13 cailloux
- Boîte 5 (B5) : 18 cailloux ($3 * 6$)

Vérification du total : $6 + 8 + 12 + 13 + 18 = 57$.

Vérification des inégalités : $6 < 8 < 12$ (OK) et $12 < 13 < 18$ (OK).

Encadré 7 : Présentation des solutions et vérification

Discussion

Ces deux « runs » montrent d'une part la capacité de Gemini (mais vraisemblablement d'autres systèmes aussi) de répondre à des requêtes de mise en forme ou de méthode.

On notera toutefois :

- La mise en forme Peano est juste accolée à des expressions algébriques sans être d'une utilité particulière.

- Dans le premier cas, 6 est explicitement déclaré la valeur minimale du contenu de la première boîte (sans explicitation du raisonnement complet). Dans le deuxième cas, cette valeur est liée à des hypothèses sur le contenu des boîtes 2 et 4.

Dans le premier cas, les autres possibilités du contenu de la première boîte sont testées. Ce n'est pas le cas dans le deuxième cas.

- Le mode de production, une marche en avant (bien que plusieurs procédures peuvent coexister avant un choix l'expression du résultat), n'a pas permis au système de supprimer la preuve inutile (et laborieuse) du nombre de cailloux supérieur à 1 dans la première boîte. Le

remarquable est qu'il s'en *souvient* (en précisant que la valeur obtenue 6 confirme ce résultat. Le déplorable est qu'il brouille la compréhension de qui veut comprendre et/ou utiliser le résultat. L'aurait-il supprimer suite à une nouvelle interrogation ? Peut-être ? Il n'empêche que cela demande de l'utilisateur un regard critique exercé que le système générateur n'a pas.

- L'exigence d'utiliser les propriétés des inégalités a plutôt alourdi le processus sans obtenir la solution obtenue en IN. Une relance ou une amélioration de l'invite permettrait peut-être d'arriver à cette solution.

Les deux derniers points soulèvent la question de la place que peut prendre ces systèmes (dans leur état actuel) dans la résolution de MWP. Tout d'abord, leur capacité à proposer un plan ou une démarche s'est passablement développée et suscite une certaine incrédulité². Par rapport à nos premiers essais, ils ont gagnés en *assurance* dans la recherche des solutions multiples. Toutefois, il semble difficile de leur faire trouver les démarches courtes de l'IN. A moins d'un travail important dans la formulation des invites. Ce qui limite évidemment leur utilité.

LOP & AF, 06.2025

Bibliographie

[1] Systèmes génératifs et résolution de problèmes mathématiques : La collection de cailloux. (<http://www.projet-ermitage.org/thema2/Chronique-cailloux.pdf>)

[2] Gabriel Poesia and Noah D. Goodman (2022). Peano: Learning Formal Mathematical Reasoning (2211.15864.pdf)

[3] Alain Favre & Luc-Olivier Pochon (2024). Systèmes génératifs et résolution de problèmes mathématiques scolaires : une chronique Le problèmes des chameaux et des dromadaires (<http://www.projet-ermitage.org/thema2/Chronique-chameaux.pdf>)

2 Et nous sommes à la recherche d'indices qui permettent de juger des procédés mis en oeuvre pour le développement des systèmes (aspect génératif général, renforcement, schémas pré-établis, distillation, etc.) et leur mise à jour (voir quelques questions posées dans [3]).