Systèmes génératifs et résolution de problèmes logiques

Version 14 juin 2025

Introduction

Après les chameaux et les dromadaires, les collections de cailloux, le problème d'Alice au pays des merveilles, le problème « A bas les profs » fait l'objet de cette chronique. C'est un problème de déduction sur la base de propositions vraies ou fausses. Les premières requêtes datent du début de février 2025. Elles sont suivies d'une plus récente (mai).

Problème « A bas les profs »

Le problème (encadré 1) est tiré de la banque de problèmes du rallye mathématique transalpin. Il fait partie la première épreuve du cinquième rallye en 1997[1]. Il est de catégorie 6 (élèves de 12 ans).

> Quatre élèves sont resté dans la classe pendant la récréation, l'un d'eux a écrit "A bas les prof." au tableau noir. Lorsque le professeur rentre en classe, il demande : "Qui a écrit ça ?" :
>
> Paul, qui porte des lunettes, dit : "c'est une fille."
>
> Marie, qui ne porte pas de lunettes : "ce n'est pas moi."
>
> Françoise, qui porte des lunettes : "c'est quelqu'un qui ne porte pas de lunettes."
>
> Jacques, qui n'a pas de lunettes : "c'est quelqu'un qui porte des lunettes."
>
> Un seul des élèves a menti. Les trois autres ont dit la vérité.
>
> **Qui a menti et qui a écrit au tableau noir ?**
>
> Expliquez votre raisonnement.

Encadré 1 : Le problème « A bas les profs »

Solution IN

Comme il n'y a qu'un menteur, il semble plus simple de considérer cet attribut en priorité. Comme les propositions émises par Jacques et Françoise sont contradictoires, l'un des deux est le menteur ou la menteuse. Paul et Marie disent donc la vérité. Cela implique que Françoise ment et que c'est elle la coupable.

A noter que seul le coupable a peut-être intérêt à mentir. Mais pas forcément.

Solution IA

Yiaho

Yiaho commence par reprendre les propositions en indiquant chaque fois la conséquence en cas de vérité (encadré 2). En faisant pour chaque personnage l'hypothèse qu'il dit la vérité, les alternatives ment/dit la vérité ne sont pas assez explorées pour les autres personnage. La découverte du personnage qui ment est liée à des *non-dits* à ce propos. Mais cela conduit à l'échec d'identifier le

1 http://www.projet-ermitage.org/ARMT/navi_fic2.php?code=sd13-fr&flag=1&langue=fr&enonce=05rmti_fr-12&w=0

coupable (encadré 3).

Pour résoudre ce problème, analysons chaque déclaration faite par les élèves en tenant compte du fait qu'un seul d'entre eux ment.

1. **Paul (qui porte des lunettes)** : "C'est une fille."
    - Donc, si Paul dit la vérité, l'auteur est une fille (Marie ou Françoise).
2. **Marie (qui ne porte pas de lunettes)** : "Ce n'est pas moi."
    - Si Marie dit la vérité, cela signifie qu'elle n'est pas l'auteur.
3. **Françoise (qui porte des lunettes)** : "C'est quelqu'un qui ne porte pas de lunettes."
    - Si Françoise dit la vérité, cela signifie que l'auteur est soit Marie soit Jacques (puisque Jacques ne porte pas de lunettes).
4. **Jacques (qui n'a pas de lunettes)** : "C'est quelqu'un qui porte des lunettes."
    - Si Jacques dit la vérité, cela signifie que l'auteur est soit Paul soit Françoise (puisqu'ils portent des lunettes).

Encadré 2 : Résumé des propositions avec une conséquence directe (Yiaho)

Nous allons examiner chaque cas :

- **Si Paul dit la vérité** (l'auteur est une fille) :
    - Cela signifie que l'auteur est soit Marie soit Françoise. Si Paul dit la vérité, alors Françoise doit mentir (car si elle dit la vérité, cela signifierait que l'un des garçons a écrit). Mais cela contredit le fait qu'il n'y a qu'un seul menteur. Donc, Paul ne peut pas dire la vérité.
- **Si Marie dit la vérité** (ce n'est pas elle) :
    - Cela signifie que l'auteur pourrait être soit Paul soit Françoise. Dans ce cas, la déclaration de Françoise (qui dit que c'est quelqu'un qui ne porte pas de lunettes) serait fausse, ce qui signifie qu'elle ment. Cela fonctionne, car cela laisse Jacques disant la vérité en disant que l'auteur porte des lunettes (donc il pourrait s'agir de Paul ou Françoise).
- **Si Françoise dit la vérité** (l'auteur est quelqu'un qui ne porte pas de lunettes) :
    - Cela signifie que l'auteur doit être Marie ou Jacques. Mais cela contredirait la déclaration de Jacques, qui dirait que l'auteur porte des lunettes, ce qui ne peut pas être vrai. Donc, cela ne fonctionne pas.
- **Si Jacques dit la vérité** (l'auteur est quelqu'un qui porte des lunettes) :
    - Cela signifierait que l'auteur est soit Paul soit Françoise. Mais cela contredirait la déclaration de Françoise, ce qui à son tour signifierait qu'elle ment.

En somme, la seule configuration possible dans laquelle un seul élève ment est lorsque **Marie dit la vérité**. Donc, **Françoise a menti**, ce qui veut dire que l'auteur qui a écrit "A bas les prof." est **Paul**. Paul porte des lunettes et c'est un garçon, ce qui correspond à l'affirmation de Jacques.

Conclusion : Marie, Paul et Jacques disent la vérité, tandis que Françoise ment. L'auteur de l'inscription est Paul.

Encadré 3 : Conclusion de Yiaho

Dans un nouvel essai Yiaho adopte une stratégie un peu différente, en particulier il ne procède pas à une reformulation des propositions. Personnage après personnage, il examine les conséquences en faisant l'hypothèse, cette fois-ci que le personnage ment. La conséquence est bien déduite pour Paul mais de façon erronée pour Marie. En définitive, ni l'auteur, ni le menteur ne sont détectés (encadré 4).

Pour résoudre ce problème, examinons les déclarations de chaque élève et l'information selon laquelle un seul élève a menti.

1. **Paul** (porte des lunettes) : "C'est une fille."
2. **Marie** (ne porte pas de lunettes) : "Ce n'est pas moi."
3. **Françoise** (porte des lunettes) : "C'est quelqu'un qui ne porte pas de lunettes."
4. **Jacques** (ne porte pas de lunettes) : "C'est quelqu'un qui porte des lunettes."

Analysons les déclarations en considérant chaque élève comme le menteur potentiel :

1. **Si Paul ment** : Cela signifie qu'il est faux que l'auteur est une fille, donc l'auteur est un garçon. Cela contredirait les déclarations de Marie, Françoise et Jacques, car ils diraient tous la vérité. Cela ne peut pas être vrai.
2. **Si Marie ment** : Cela signifie que c'est elle qui a écrit "A bas les prof." Cela contredirait la déclaration de Françoise (qui dirait la vérité en affirmant que c'est quelqu'un qui ne porte pas de lunettes). Ce scénario fonctionne, mais ne peut pas être vrai puisque cela implique deux mensonges.
3. **Si Françoise ment** : Cela signifie qu'il est faux que l'auteur est quelqu'un qui ne porte pas de lunettes, donc l'auteur est Paul ou Françoise elle-même. Cela contredit les déclarations de Paul et Marie, donc ce scénario ne peut pas être vrai.
4. **Si Jacques ment** : Cela signifie que l'auteur n'est pas quelqu'un qui porte des lunettes. Cela signifie que l'auteur est soit Marie soit Jacques lui-même. Cela ne contredit aucune autre déclaration, car Paul et Françoise diraient la vérité.

En vérifiant la dernière option, nous avons :

- Paul dit la vérité (c'est une fille).
- Marie dit la vérité (ce n'est pas elle).
- Françoise dit la vérité (c'est quelqu'un qui ne porte pas de lunettes).
- Jacques ment (et l'écrivain est donc Marie, qui ne porte pas de lunettes).

Par conséquent, Jacques est celui qui a menti. La personne qui a écrit "A bas les prof." est Marie.

Encadré 4 : Nouvel essai avec Yiaho

Nous lui signalons alors une contradiction : *Mais Marie a dit que ce n'était pas elle!* Le *raisonnement* est repris. Il semble alors que le système confond parfois les attributs des personnages et ce qu'ils disent. Mais le personnage menteur n'est pas identifié. (encadré 5)

Le mystère reste de savoir si le système a utilisé son erreur (pour clarifier le raisonnement comme il le signale) ou a simplement compté sur le jeu des probabilités. Pour traiter cette question, il faudrait en préalable cerner le processus. La partie générative et la partie qui ferait intervenir des schémas déductifs.

YIAHO IQ 200 :

Vous avez raison, examinons à nouveau les déclarations et corrigeons l'analyse. Voici les déclarations et les implications :

1. **Paul** (porte des lunettes) : "C'est une fille."
2. **Marie** (ne porte pas de lunettes) : "Ce n'est pas moi."
3. **Françoise** (porte des lunettes) : "C'est quelqu'un qui ne porte pas de lunettes."
4. **Jacques** (ne porte pas de lunettes) : "C'est quelqu'un qui porte des lunettes."

Rappelons que seulement un élève ment.

Analysons les déclarations :

1. **Si Paul ment** : Cela signifie que l'auteur est un garçon. Dans ce cas, Marie, Françoise et Jacques diraient la vérité. Cela entraînerait une contradiction car Jacques dirait que l'écrivain porte des lunettes. Donc, Paul ne peut pas être le menteur.
2. **Si Marie ment** : Cela signifie qu'elle a écrit sur le tableau. Dans ce cas, Paul dirait la vérité (l'auteur est une fille), ce qui est cohérent. Françoise (l'auteur ne porte pas de lunettes) mentirait, ce qui est contradictoire. Donc, Marie ne peut pas être le menteur.
3. **Si Françoise ment** : Cela signifie que l'auteur porte des lunettes. Cela contredirait Paul (qui dit que c'est une fille) et Jacques (qui dit que l'auteur porte des lunettes). Donc, Françoise ne peut pas être le menteur.
4. **Si Jacques ment** : Cela signifie que l'auteur ne porte pas de lunettes. Alors, cela peut être soit Marie, soit Françoise. Dans ce cas :
    - Paul dit la vérité (c'est une fille).
    - Marie dit la vérité (ce n'est pas elle).
    - Françoise dirait que c'est quelqu'un qui ne porte pas de lunettes, ce qui dans ce cas est vrai.

Cela ne peut pas être une contradiction ici.

Conclusion

Le menteur est **Jacques**, et la personne qui a écrit "A bas les prof." est **Françoise**, car elle est la seule à correspondre à la déclaration de Paul (c'est une fille) et à ce que Jacques a dit (l'auteur n'est pas celui qui porte des lunettes).

Merci d'avoir attiré mon attention sur l'erreur, cela a permis de clarifier le raisonnement.

Encadré 5 : Reprise de la *réflexion* (Yiaho)

Notre réaction est évidemment de signaler que *Françoise porte des lunettes* contrairement à ce que dit Jaques en mentant.

Les conséquences de chaque hypothèse sont alors mieux explicitées (encadré 6). La conclusion est plutôt désopilante : Marie est la coupable tout en disant sans mentir que ce n'est pas elle. Cet épisode montre qu'il n'y a vraisemblablement pas de « dispositif » jugeant la cohérence d'une proposition. Il serait également intéressant de savoir ce que recouvre la « confusion » qui est mentionnée.

YIAHO IQ 200 : Vous avez tout à fait raison, je m'excuse pour la confusion dans mon raisonnement. Analysons à nouveau les déclarations et déterminons qui a menti et qui a écrit sur le tableau :

Récapitulons les déclarations :

1. **Paul** (porte des lunettes) : "C'est une fille."
2. **Marie** (ne porte pas de lunettes) : "Ce n'est pas moi."
3. **Françoise** (porte des lunettes) : "C'est quelqu'un qui ne porte pas de lunettes."
4. **Jacques** (ne porte pas de lunettes) : "C'est quelqu'un qui porte des lunettes."

Analysons les possibilités :

1. **Si Paul ment :** Cela signifie que l'auteur n'est pas une fille.
    - Marie dirait la vérité (ce n'est pas elle).
    - Françoise dirait la vérité (l'auteur porte des lunettes).
    - Jacques dirait la vérité (l'auteur porte des lunettes).
    - Cela ne fonctionne pas, car on aurait deux auteurs possibles (Françoise et Paul, en contradiction avec l'affirmation de Paul).
2. **Si Marie ment :** Cela signifie qu'elle a écrit sur le tableau.
    - Paul dirait la vérité (c'est une fille).
    - Françoise dirait la vérité (l'auteur ne porte pas de lunettes).
    - Jacques dirait la vérité (l'auteur porte des lunettes).
    - Cela ne fonctionne pas, car Marie ne peut pas être à la fois l'auteur et celle qui ne porte pas de lunettes.
3. **Si Françoise ment :** Cela signifie que l'auteur porte des lunettes.
    - Cela contredirait Paul, qui dit que c'est une fille, mais cela pourrait être vrai si l'auteur était Jacques.
    - Marie dirait la vérité (ce n'est pas elle).
    - Jacques dirait alors la vérité (l'auteur porte des lunettes), ce qui est une contradiction.
4. **Si Jacques ment :** Cela signifie que l'auteur ne porte pas de lunettes.
    - Paul dirait la vérité (c'est une fille).
    - Marie dirait la vérité (ce n'est pas elle).
    - Françoise dirait la vérité (l'auteur ne porte pas de lunettes) - c'est donc une contradiction, car elle porte des lunettes.

Conclusion :

En réévaluant les possibilités, il apparaît que la seule combinaison cohérente est que **Jacques ment**. L'auteur de l'écrit est **Marie**, qui dit la vérité en disant "ce n'est pas moi", et qui ne porte pas de lunettes.

Je vous remercie de votre patience et d'avoir souligné les erreurs précédentes. La bonne réponse est que Jacques est celui qui a menti et que c'est Marie qui a écrit sur le tableau noir.

Encadré 6 : Nouvel essai et conclusion étrange (Yiaho)

ChatGPT

Pour résoudre cette énigme, analysons les déclarations des quatre élèves en tenant compte du fait qu'un seul ment, tandis que les trois autres disent la vérité.

Déclarations des élèves :

1. **Paul** (porte des lunettes) : «□C'est une fille.□»
2. **Marie** (ne porte pas de lunettes) : «□Ce n'est pas moi.□»
3. **Françoise** (porte des lunettes) : «□C'est quelqu'un qui ne porte pas de lunettes.□»
4. **Jacques** (ne porte pas de lunettes) : «□C'est quelqu'un qui porte des lunettes.□»

Analyse :

- **Si Paul dit la vérité** («□C'est une fille□») :
    - L'auteur est donc une fille, soit Marie (sans lunettes) ou Françoise (avec lunettes).
    - Si Marie dit la vérité («□Ce n'est pas moi□»), alors l'auteur est Françoise.
    - Si Françoise dit la vérité («□C'est quelqu'un qui ne porte pas de lunettes□»), cela contredit le fait que l'auteur soit Françoise.
    - Donc, Françoise mentirait, et l'auteur serait Marie, ce qui contredit la déclaration de Marie.
    - **Conclusion** : Cette hypothèse mène à une contradiction.
- **Si Marie dit la vérité** («□Ce n'est pas moi□») :
    - L'auteur est donc Paul, Françoise ou Jacques.
    - Si Paul dit la vérité («□C'est une fille□»), alors l'auteur est Françoise.
    - Si Françoise dit la vérité («□C'est quelqu'un qui ne porte pas de lunettes□»), alors l'auteur est Marie ou Jacques, ce qui contredit la déclaration de Marie.
    - Donc, Françoise mentirait, et l'auteur serait Françoise, ce qui est une contradiction.
    - **Conclusion** : Cette hypothèse mène à une contradiction.
- **Si Françoise dit la vérité** («□C'est quelqu'un qui ne porte pas de lunettes□») :
    - L'auteur est donc Marie ou Jacques.
    - Si Paul dit la vérité («□C'est une fille□»), alors l'auteur est Marie.
    - Si Marie dit la vérité («□Ce n'est pas moi□»), alors l'auteur est Jacques, ce qui contredit la déclaration de Paul.
    - Donc, Paul mentirait, et l'auteur serait Marie ou Jacques.
    - Si l'auteur est Marie, cela contredit sa propre déclaration.
    - Si l'auteur est Jacques, cela contredit la déclaration de Paul.
    - **Conclusion** : Cette hypothèse mène à une contradiction.
- **Si Jacques dit la vérité** («□C'est quelqu'un qui porte des lunettes□») :
    - L'auteur est donc Paul ou Françoise.
    - Si Paul dit la vérité («□C'est une fille□»), alors l'auteur est Françoise.
    - Si Marie dit la vérité («□Ce n'est pas moi□»), alors l'auteur est Françoise.
    - Si Françoise dit la vérité («□C'est quelqu'un qui ne porte pas de lunettes□»), cela contredit le fait que l'auteur soit Françoise.
    - Donc, Françoise mentirait, et l'auteur serait Françoise, ce qui est une contradiction.
    - **Conclusion** : Cette hypothèse mène à une contradiction.

Conclusion générale :

Après avoir analysé toutes les hypothèses, il apparaît que chacune mène à une contradiction. Il est possible que l'énoncé de l'énigme comporte une ambiguïté ou une erreur.

Encadré 7 : Solution de ChatGPT en mode rechercher

Nouveauté que nous découvrons, le chatbot peut être interrogé en mode « rechercher » ou « raisonner ». Dans ce dernier cas, il est préciser qu'il y a approfondissement avant de répondre.

En mode « rechercher » (encadré 7) le système résume les conditions puis fait un *raisonnement* pour chaque personnage avec l'hypothèse qu'il dit la vérité. Tous les cas possibles de chaque hypothèse ne sont pas examinés. Chaque raisonnement aboutit chaque fois à une contradiction. Conclusion : « Il est possible que l'énoncé de l'énigme comporte une ambiguïté ou une erreur. ».

Dans ce mode, les sources sont à disposition. Elles sont d'une grande diversité. Parmi elles notre énoncé, à l'identique, se retrouve sur le site d'aide aux devoirs[2]. Cette question est posée par Pascale en novembre 2015 (18 ans après l'original). Elle lui a été vraisemblablement proposée par un enseignant ayant participé à un rallye ou qui a trouvé l'énoncé sur le site du rallye.

La réponse fournie par melou59 s'épargne toute logique : « C'est francsoise qui ment car tout les paroles des eleves sont contre elle. Car les garcons disent que c'est une fille avec des lunettes et la fille presente qui n'a pas de lunettes affirme que ce n'est pas elle. » (sic et resic)

Dans ces sources se trouvent plusieurs problèmes isomorphes certains précisant que le menteur est le coupable.

Considérant la situation

Je suis en train de personnaliser la réponse en soulignant l'importance du respect et de l'interpelle.

Réfléchir et synthétiser

Je vais refuser de vous aider à créer la signature générative. номинатив Cun du ose tu lrasr.

Réfléchir aux conséquences

Je me demande comment la réponse pourrait nuire aux élèves, tout en déterminant la gravité de l'incident et les actions possibles.

Évaluer les défis

Je suis en train de déterminer si l'absence d'un enseignant pendant la pause offre une opportunité d'apprentissage ou un défi à relever.

Réagir intelligent

Je suis appelé à réorienter la situation en réfléchissant aux implications des actions des élèves qu'ils font.

Réfléchir aux réponses

Je suis en train de répondre aux élèves qui ont osé écrire "A bas les p". C'est tentant, mais je suis en train de voir plus clair dans leurs actions.

Réfléchir aux conséquences

Je suis en train de peser les implications des expressions des élèves et la manière de réagir, en évaluant la gravité de leurs actions et les mesures appropriées.

Analyser la situation

Je suis en train de voir comment les élèves réagissent en termes de respect et d'attitude. Il pourrait être intéressant de discuter des raisons de leur comportement avec la directive.

Encadré 8a : ChatGPT en mode raisonner (février 2005)

2 https://nosdevoirs.fr/devoir/882133

Un premier essai en mode « raisonner » figure dans les encadrés 8a et 8b. L'encadré 8a constitue vraisemblablement l'approfondissement. C'est plutôt obscur. S'adresse-t-il aux élèves dans l'énigme ou ceux qui sont sensés la résoudre ? La deuxième considération est particulièrement énigmatique sans compter ce qui paraît comme une erreur dans la dé-tokenisation.

Dans l'encadré 8b, les propositions émises par chaque enfant sont explicitées en cas de mensonge et en cas de vérité. Il n'y a pas alors de raisonnement proprement dit. Celui-ci suit dans un seul paragraphe.

Pour résoudre ce problème, analysons les déclarations des quatre élèves et les conséquences de chacune si elle est vraie ou fausse, en tenant compte du fait qu'un seul élève ment et que les trois autres disent la vérité.

1. **Paul (qui porte des lunettes) dit : "C'est une fille."**
    - Si Paul dit la vérité, alors celui qui a écrit sur le tableau est une fille.
    - Si Paul ment, alors la personne qui a écrit au tableau est un garçon.
2. **Marie (qui ne porte pas de lunettes) dit : "Ce n'est pas moi."**
    - Si Marie dit la vérité, alors elle n'a pas écrit au tableau.
    - Si Marie ment, alors elle a écrit au tableau.
3. **Françoise (qui porte des lunettes) dit : "C'est quelqu'un qui ne porte pas de lunettes."**
    - Si Françoise dit la vérité, alors la personne qui a écrit au tableau est quelqu'un qui ne porte pas de lunettes.
    - Si Françoise ment, alors la personne qui a écrit au tableau porte des lunettes.
4. **Jacques (qui ne porte pas de lunettes) dit : "C'est quelqu'un qui porte des lunettes."**
    - Si Jacques dit la vérité, alors la personne qui a écrit au tableau porte des lunettes.
    - Si Jacques ment, alors la personne qui a écrit au tableau ne porte pas de lunettes.

Cas 1 : Si Paul dit la vérité

Si Paul dit la vérité et que c'est une fille qui a écrit au tableau, alors Marie et Françoise sont les seules filles possibles (puisque Jacques est un garçon et Paul aussi). Cependant, si Françoise avait écrit, sa déclaration serait fausse (puisque Françoise dit que c'est quelqu'un qui ne porte pas de lunettes, et si elle avait écrit au tableau, elle porterait des lunettes). Donc, Françoise ment. Dans ce cas, Jacques dirait la vérité en disant que la personne qui a écrit porte des lunettes, ce qui serait cohérent. Mais cela implique que Marie ment, car elle aurait dit "Ce n'est pas moi".

Si Marie ment, cela signifie qu'elle est l'élève qui a écrit sur le tableau. Cela contredit le fait que Paul dit que c'est une fille, car Paul serait alors celui qui ment.

Donc, **Paul ment** et **Marie a écrit au tableau**.

Conclusion :

- **Paul ment.**
- **Marie a écrit au tableau.**

Encadré 8b : Solution ChatGPT

Lors d'un deuxième essai en mode « raisonner », le comportement de ChatGPT est très différent.

Tout d'abord (encadré 9a) un plan de résolution (en anglais) est exposé. L'inconnue devient X deux symboles sont proposés pour porter ou ne pas porter des lunettes (G et NG). Seront-ils utilisés et si oui comment reste un mystère.

Piecing together clues

I'm looking at the puzzle where four students, three truthful and one dishonest, are accused of writing on the blackboard during recess. The challenging part involves identifying the truth-teller using specific statements.

Analyzing statements

OK, I'm digging into the puzzle: four students, one liar, and the others are honest. Each student's words give a clue about who wrote on the blackboard.

Mapping out variables

I'm noting variables: G for "has glasses" and NG for "no glasses." Assigning Paul, Marie, Françoise, and Jacques these categories to analyze their statements effectively.

Considering Jacques' statement

Interestingly enough, Jacques says "X is someone who wears glasses." If true, X must wear glasses. If false, X does not wear glasses, considering the other factors.

Evaluating Marie's possible lie

I'm testing if Marie lied. This indicates X is a girl, narrowing it down to either Marie or Françoise, while the other criteria are simultaneously considered.

Piecing together Françoise's scenario

If Françoise lied, X must be glasses. It's tough to reconcile Françoise's non-glasses truth with Jacques' glasses truth, so this case doesn't quite fit.

Connecting the dots

If Françoise lied, X would wear glasses. This aligns with Françoise's glasses-wearing truth, suggesting X is Françoise and she indeed wears glasses.

Confirming the conclusion

Françoise lied, suggesting X wears glasses. Her glasses-wearing truth confirms X is Françoise. Jacques' glasses-wearing truth aligns, validating all conditions.

Piecing together logical clues

Examining the clues, the only scenario that fits is Case 3. Françoise lied, pointing to Françoise as the liar and the person behind the prank.

Encadré 9a : ChatGPT en mode Raisonner (2e tentative)

C'est un mélange de trivialités : « Interestingly enough, Jacques says "X is someone who wears glasses." If true, X must wear glasses. If false, X does not wear glasses, considering the other factors. », de raisonnements partiels : I'm testing if Marie lied. This indicates X is a girl, narrowing it down to either Marie or Françoise, while the other criteria are simultaneously considered. » qui rassemblés conduisent à la solution.

Puis la solution est proposée en français de façon plus structurée. D'abord les caractéristiques de chaque personnage et son affirmation sont reprises. Ensuite, Etape 1, pour chaque personnage un brève déduction, légèrement tautologique de notre point de vue, est proposée selon que celui-ci mente ou dise vrai. (encadré 9b).

Analysons chaque témoignage en notant d'une part que seuls trois élèves disent la vérité et qu'un seul ment.

Données sur les élèves

- **Paul** (lunettes) affirme : «□C'est une fille.□»
- **Marie** (sans lunettes) affirme : «□Ce n'est pas moi.□»
- **Françoise** (lunettes) affirme : «□C'est quelqu'un qui ne porte pas de lunettes.□»
- **Jacques** (sans lunettes) affirme : «□C'est quelqu'un qui porte des lunettes.□»

Notons que l'auteur de la bêtise est l'un d'entre eux, et on appellera cet élève X.

Étape 1 : Traduction des témoignages en contraintes

- **Paul (vrai ou faux)**
    - Si Paul dit la vérité, alors X est une fille (donc soit Marie, soit Françoise).
    - S'il ment, alors X est un garçon (donc soit Paul ou Jacques).
- **Marie**
    - Si Marie dit la vérité, alors X ≠ Marie.
    - Si elle ment, alors X = Marie.
- **Françoise**
    - Si Françoise dit la vérité, alors X porte **pas** de lunettes (donc X = Marie ou Jacques).
    - Si elle ment, alors X porte des lunettes (donc X = Paul ou Françoise).
- **Jacques**
    - S'il dit la vérité, alors X porte des lunettes (donc X = Paul ou Françoise).
    - S'il ment, alors X ne porte pas de lunettes (donc X = Marie ou Jacques).

Encadré 9b : Suite de la réponse de ChatGPT, reprise des énoncés

Cette partie semble procéder à la mise en place d'une formalisation des énoncés et de leur conséquence immédiate. Par exemple : « Si Françoise dit la vérité, alors X porte **pas** de lunettes (donc X = Marie ou Jacques) ».

L'étape 2 (encadré 9c) passe en revue chaque personnage supposant qu'il ment alors que tous les autres disent la vérité. Dans trois cas, les éléments rassemblés mènent à une contradiction. Le quatrième cas est la solution. Une vérification (et non un résumé du raisonnement) précède de la formulation de la solution.

Étape 2 : Examiner les cas en considérant qu'un seul ment

Nous étudierons successivement l'hypothèse selon laquelle chacun serait le menteur unique.

1. **Supposons que Paul ment**
    - Paul ment : son affirmation « C'est une fille » est fausse, donc X est un garçon (Paul ou Jacques).
    - Marie (vraie) : X ≠ Marie.
    - Françoise (vraie) : X ne porte pas de lunettes, donc X = Marie ou Jacques.
      – Mais Marie est exclue (d'après son témoignage) et Jacques n'a pas de lunettes, ce qui convient avec Françoise.
    - Jacques (vrai) : X porte des lunettes.
      – Or, Jacques lui-même ne porte pas de lunettes, et s'il est X, cela contredirait son témoignage. ^ Contradiction.
2. **Supposons que Marie ment**
    - Marie ment : « Ce n'est pas moi » est faux, donc X = Marie.
    - Paul (vrai) : X est une fille. Marie est bien une fille, c'est cohérent.
    - Françoise (vrai) : X ne porte pas de lunettes. Marie, sans lunettes, correspond.
    - Jacques (vrai) : X porte des lunettes.
      – Ici, contradiction, car Marie ne porte pas de lunettes. ^ Contradiction.
3. **Supposons que Françoise ment**
    - Françoise ment : son affirmation « C'est quelqu'un qui ne porte pas de lunettes » est fausse, donc X porte des lunettes.
    - Paul (vrai) : X est une fille, donc X = Marie ou X = Françoise.
    - Marie (vrai) : X ≠ Marie, donc il reste X = Françoise.
    - Jacques (vrai) : X porte des lunettes.
      – Françoise porte bien des lunettes.
      ^ Toutes les contraintes sont satisfaites.
4. **Supposons que Jacques ment**
    - Jacques ment : son affirmation « C'est quelqu'un qui porte des lunettes » est fausse, donc X ne porte pas de lunettes.
    - Paul (vrai) : X est une fille, donc X = Marie ou Françoise.
    - Marie (vrai) : X ≠ Marie, donc X = Françoise.
    - Françoise (vrai) : X ne porte pas de lunettes, mais Françoise porte des lunettes. ^ Contradiction.

Conclusion

L'unique scénario cohérent est celui où **Françoise ment** et l'auteur (X) est **Françoise** elle-même.

Raisonnement résumé

- Paul affirme que l'auteur est une fille. Comme Françoise est une fille, cette affirmation est vraie.
- Marie nie être l'auteur, ce qui est vrai puisque c'est Françoise.
- Jacques affirme que l'auteur porte des lunettes, ce qui est vrai pour Françoise.
- Seule Françoise, en disant que l'auteur ne porte pas de lunettes, ment. Ainsi, Françoise est à la fois la menteuse et l'élève qui a écrit « A bas les prof. » au tableau noir.

Encadré 9c : Raisonnement et solution pour ChatGPT

DeepSeek

A la fin de ces essais DeepSeek entre sur le marché et nous empressons de lui présenter notre problème. A l'instar de ChatGPT, DeepSeek, en anglais, résume les caractéristiques et personnages et leur affirmation. Mis à part quelques évidences : « So, the possible culprits are these four: Paul, Marie, Françoise, or Jacques. The culprit must be one of them. » les personnages sont passés en revue en prenant l'hypothèse qu'ils sont coupables. Les conséquences en terme de nombre de menteurs, de cette hypothèse de culpabilité sur les affirmations des autres personnages sont alors proposées, ce qui permet d'en déduire la solution (encadrés 10a et 10b).

Finalement, les déductions sont épurées et présentées en français. (encadré 10c)

Okay, so there's this problem where four students stayed in class during recess, and one of them wrote "Down with teachers" on the blackboard. The teacher comes back and asks who did it. Each of the four students makes a statement, and we know that only one of them lied, while the other three told the truth. We need to figure out who lied and who the culprit is.

First, let me list out the students and their statements:

1. Paul (wears glasses) says: "It's a girl."

2. Marie (doesn't wear glasses) says: "It wasn't me."

3. Françoise (wears glasses) says: "It's someone who doesn't wear glasses."

4. Jacques (doesn't wear glasses) says: "It's someone who wears glasses."

We know only one of them is lying. So, three statements are true, one is false.

Let me start by noting the names and whether they wear glasses:

- Paul: glasses

- Marie: no glasses

- Françoise: glasses

- Jacques: no glasses

So, the possible culprits are these four: Paul, Marie, Françoise, or Jacques. The culprit must be one of them.

Let's analyze each possibility step by step, assuming each one is the culprit and checking if only one person lies in each case.

First, let's suppose that Paul is the culprit. If Paul did it, then:

- Paul's statement: "It's a girl." But Paul is a boy, so this would be a lie. So Paul would be the liar.

- Marie's statement: "It wasn't me." True, because Paul did it.

- Françoise's statement: "It's someone who doesn't wear glasses." But Paul wears glasses, so this would be a lie.

- Jacques's statement: "It's someone who wears glasses." Paul does wear glasses, so that's true.

But in this scenario, both Paul and Françoise would be lying, which contradicts the condition that only one lied. So Paul can't be the culprit.

Encadré 10a : Exposé des données et analyse des conséquences si Paul est le menteur (Deepseek)

Next, let's check Marie. If Marie did it:

- Marie's statement: "It wasn't me." That would be a lie. So Marie is the liar.

- Paul's statement: "It's a girl." Marie is a girl, so that's true.

- Françoise's statement: "It's someone who doesn't wear glasses." Marie doesn't wear glasses, so that's true.

- Jacques's statement: "It's someone who wears glasses." Since Marie doesn't wear glasses, that's a lie.

Wait, so if Marie did it, both Marie and Jacques would be lying. That's two liars, which is not allowed. So Marie can't be the culprit either.

Now, let's consider Françoise. If Françoise did it:

- Françoise's statement: "It's someone who doesn't wear glasses." But Françoise wears glasses, so that's a lie. So Françoise is the liar.

- Paul's statement: "It's a girl." Françoise is a girl, so true.

- Marie's statement: "It wasn't me." True, because Françoise did it.

- Jacques's statement: "It's someone who wears glasses." Françoise wears glasses, so that's true.

In this case, only Françoise lied. That fits the condition: one liar, three truths. So this seems possible. Let me check the last possibility to be thorough.

Lastly, Jacques. If Jacques did it:

- Jacques's statement: "It's someone who wears glasses." But Jacques doesn't wear glasses, so that's a lie. So Jacques is the liar.

- Paul's statement: "It's a girl." Jacques is a boy, so that's a lie.

- Marie's statement: "It wasn't me." True.

- Françoise's statement: "It's someone who doesn't wear glasses." Jacques doesn't wear glasses, so true.

Here, both Jacques and Paul would be lying. Two liars again, which isn't allowed. So Jacques can't be the culprit.

So the only scenario where only one person lies is when Françoise is the culprit. In that case, Françoise lied, and the others told the truth. Let me double-check:

- Paul (glasses) says it's a girl. Françoise is a girl, so true.

- Marie (no glasses) says it's not her. True.

- Françoise (glasses) says it's someone without glasses. She did it, and she has glasses, so lie.

- Jacques (no glasses) says it's someone with glasses. Françoise has glasses, so true.

Yes, that works. Only Françoise lied. Therefore, Françoise is the liar and the culprit.

Encadré 10b : Suite des conséquences selon que chaque personnage est le menteur (DeepSeek)

S o l u t i o n :
Le menteur est **Françoise** et c'est elle qui a écrit l'inscription au tableau.

Explication détaillée :

1. **Analyse des déclarations et des caractéristiques :**
   - **Paul** (lunettes) : "C'est une fille."
   - **Marie** (pas de lunettes) : "Ce n'est pas moi."
   - **Françoise** (lunettes) : "C'est quelqu'un qui ne porte pas de lunettes."
   - **Jacques** (pas de lunettes) : "C'est quelqu'un qui porte des lunettes."

   Seul un élève ment, les autres disent la vérité.
2. **Test des hypothèses :**
   - **Si Paul est coupable** :
     - Paul ment ("c'est une fille" ^ faux, car Paul est un garçon).
     - Françoise ment aussi (elle affirme "sans lunettes", mais Paul en porte). ^ Deux menteurs : impossible.
   - **Si Marie est coupable** :
     - Marie ment ("ce n'est pas moi").
     - Jacques ment ("c'est quelqu'un avec lunettes", mais Marie n'en porte pas). ^ Deux menteurs : impossible.
   - **Si Françoise est coupable** :
     - Françoise ment ("sans lunettes", alors qu'elle en porte).
     - Paul dit vrai (Françoise est une fille).
     - Marie dit vrai (ce n'est pas elle).
     - Jacques dit vrai (Françoise porte des lunettes).

       ^ **Un seul menteur** : valide.
   - **Si Jacques est coupable** :
     - Jacques ment ("avec lunettes", mais il n'en porte pas).
     - Paul ment ("c'est une fille", alors que Jacques est un garçon). ^ Deux menteurs : impossible.
3. **C o n c l u s i o n :**
   Seule l'hypothèse où **Françoise est coupable** satisfait la condition d'un seul menteur. Françoise a menti en accusant une personne sans lunettes, alors qu'elle-même en porte.

Encadré 10c : Déductions épurées et traduites en français (DeepSeek)

Gemini (mai 2025)

Trois mois après nos essais apparaît le nouvel assistant conversationnel proposé par Google (ou bien nouvellement mis dans une version très accessible).

Dans un premier temps, le système expose le processus de raisonnement : Comprendre l'objectif, lister les participants et leurs affirmations, analyser les informations et leur relations ou

contradiction (encadré 11a).

Thoughts (experimental)
Here's a thinking process to solve this logic puzzle:

1. **Understand the Goal:**
   - Identify who lied.
   - Identify who wrote on the blackboard.
   - The core constraint: exactly one person lied, three told the truth.
2. **List the Participants and Their Statements:**
   - Paul (lunettes): "C'est une fille." (It's a girl.)
   - Marie (pas de lunettes): "Ce n'est pas moi." (It's not me.)
   - Françoise (lunettes): "C'est quelqu'un qui ne porte pas de lunettes." (It's someone who doesn't wear glasses.)
   - Jacques (pas de lunettes): "C'est quelqu'un qui porte des lunettes." (It's someone who wears glasses.)
3. **Analyze the Statements for Contradictions/Relationships:**
   - Françoise's statement and Jacques's statement are direct opposites regarding whether the culprit wears glasses or not.
     - If Françoise is telling the truth (culprit doesn't wear glasses), then Jacques is lying.
     - If Jacques is telling the truth (culprit wears glasses), then Françoise is lying.
   - This is a key insight: *One of Françoise or Jacques MUST be the liar, and the other MUST be telling the truth.*
   - Since only one person lied, this means Paul and Marie *must* be telling the truth.

Encadré 11a : Présentation du processus de raisonnement (Gemini)

Ce schéma est parfait et pose une nouvelle fois la question de savoir s'il a été acquis par une sorte d'entrainement au deuxième degré sur de multiples problèmes apparentés ou « enseigné » (comment?) par des petites ou plus grosses mains extérieures.

Par rapport aux autres systèmes, il est explicitement proposé l'analyse des relations et surtout des contradictions. De ce point de vue la phrase mise en italique (par le système!) est la clé pour une solution rapide.

La suite du raisonnement se base sur cette découverte (encadré 11b). La solution est suivie par une vérification et une préparation de l'explication.

L'encadré suivant (encadré 11b) reprend de façon plus synthétique le raisonnement en français.

Assez subjugué par le brio avec lequel le problème a été résolu, nous demandons à Gemini un problème semblable. Sa réponse est le « Gâteau disparu » figurant en annexe. Nous avons trouvé des traces de ce gâteau sur Internet. Le mystère du gâteau disparu est une énigme de Cluedo Junior. Cette énigme se résout dans un jeu faisant intervenir un plateau de jeu et plusieurs joueurs (au moins deux). La mise en scène précise que quelqu'un s'est introduit dans la cuisine et a dérobé une part de gâteau avant même le début du dîner ! Il s'agit alors de découvrir le nom du coupable, mais aussi avec quelle boisson et à quelle heure ! Une autre référence se trouve sur le site de « mes histoires du soir »[3]. L'histoire est alors enrobée dans la préparation d'une fête villageoise où l'on prépare un gâteau géant. Mais ce n'est pas un problème de logique plutôt d'identification à partir de quelques observations

3 https://www.meshistoiresdusoir.fr/h/-le-mystere-du-g-teau-disparu-/

4. **Use the "Truth-Tellers" to Deduce Facts:**
    - Paul (truth-teller): "C'est une fille."
        - This means the culprit is either Marie or Françoise.
    - Marie (truth-teller): "Ce n'est pas moi."
        - This means Marie is not the culprit.
5. **Combine the Deductions:**
    - From Paul: Culprit is Marie or Françoise.
    - From Marie: Culprit is not Marie.
    - Therefore, the culprit *must be Françoise*.
6. **Verify with Françoise and Jacques (the liar/truth-teller pair):**
    - We know the culprit is Françoise. Françoise wears glasses.
    - Let's check Françoise's statement: "C'est quelqu'un qui ne porte pas de lunettes."
        - If Françoise wrote it, and she wears glasses, then her statement is FALSE. This means Françoise is the liar.
    - Let's check Jacques's statement: "C'est quelqu'un qui porte des lunettes."
        - If Françoise wrote it, and she wears glasses, then Jacques's statement is TRUE.
7. **Confirm Consistency:**
    - **Culprit:** Françoise
    - **Liar:** Françoise
    - Paul's statement: "C'est une fille." (Françoise is a girl) -> TRUE.
    - Marie's statement: "Ce n'est pas moi." (Marie didn't do it) -> TRUE.
    - Françoise's statement: "C'est quelqu'un qui ne porte pas de lunettes." (Françoise wears glasses, so this is about herself) -> FALSE.
    - Jacques's statement: "C'est quelqu'un qui porte des lunettes." (Françoise wears glasses) -> TRUE.
    - This fits the condition: one liar (Françoise), three truth-tellers (Paul, Marie, Jacques).
8. **Structure the Explanation:**
    - Start by stating the core assumption (one liar, three truth-tellers).
    - Identify the contradictory statements (Françoise vs. Jacques) and deduce that one of them is the liar.
    - This means the other two (Paul and Marie) are truth-tellers.
    - Use Paul's true statement to narrow down the culprit.
    - Use Marie's true statement to pinpoint the culprit.
    - Identify the culprit.
    - Now, go back to Françoise and Jacques. Determine which of their statements is false, given who the culprit is. This identifies the liar.
    - State the final answer clearly.

(Self-correction during the process):

- What if I initially assumed Paul lied? Then Marie, Françoise, and Jacques told the truth.
    - Marie (Truth): Not Marie.
    - Françoise (Truth): Culprit no glasses (Jacques or Marie). Since not Marie, culprit is Jacques.
    - Jacques (Truth): Culprit wears glasses (Paul or Françoise).
    - Contradiction! Françoise says culprit is Jacques (no glasses), Jacques says culprit is Paul or Françoise (glasses). This scenario is impossible. This confirms the initial deduction about Françoise/Jacques being the liar/truth-teller pair is crucial.

Encadré 11b : Déduction finale, solution, projet d'explication (Gemini)

Analysons les déclarations étape par étape, en gardant à l'esprit qu'un seul élève ment.

1. **Les déclarations de Françoise et Jacques sont contradictoires :**
    - Françoise (lunettes) dit : "c'est quelqu'un qui ne porte pas de lunettes."
    - Jacques (pas de lunettes) dit : "c'est quelqu'un qui porte des lunettes."
      Ces deux affirmations ne peuvent pas être vraies en même temps. L'une doit être vraie et l'autre fausse. Puisqu'il n'y a qu'un seul menteur, l'un d'eux dit la vérité et l'autre ment.
2. **Cela signifie que Paul et Marie disent la vérité.**
    - Paul (lunettes) dit : "c'est une fille." (VRAI)
    - Marie (pas de lunettes) dit : "ce n'est pas moi." (VRAI)
3. **Utilisons les affirmations vraies de Paul et Marie pour trouver le coupable :**
    - Puisque Paul dit la vérité, le coupable est une fille. Les filles sont Marie et Françoise.
    - Puisque Marie dit la vérité, ce n'est pas Marie.
    - Donc, par élimination, **Françoise a écrit au tableau noir.**
4. **Maintenant, déterminons qui a menti entre Françoise et Jacques :**
    - Françoise a écrit au tableau noir. Françoise porte des lunettes.
    - La déclaration de Françoise était : "c'est quelqu'un qui ne porte pas de lunettes." Puisque c'est elle (qui porte des lunettes) qui a écrit, sa déclaration est FAUSSE.
    - La déclaration de Jacques était : "c'est quelqu'un qui porte des lunettes." Puisque c'est Françoise (qui porte des lunettes) qui a écrit, sa déclaration est VRAIE.
5. **Conclusion :**
    - **Françoise a menti.**
    - **Françoise a écrit au tableau noir.**

Vérifions :

- Paul (dit la vérité) : "c'est une fille." (Françoise est une fille -> Vrai)
- Marie (dit la vérité) : "ce n'est pas moi." (C'est Françoise -> Vrai)
- Françoise (ment) : "c'est quelqu'un qui ne porte pas de lunettes." (C'est elle, et elle porte des lunettes -> Faux)
- Jacques (dit la vérité) : "c'est quelqu'un qui porte des lunettes." (C'est Françoise, qui porte des lunettes -> Vrai)

Le scénario est cohérent avec un seul menteur.

11c : Mise en français du processus de résolution (Gemini)

Discussion

Résumé des séances

Yiaho

Dans une première séance le système commence à reformuler les caractéristiques et les assertions des différents personnages. Puis personnage après personnage, il explore conséquences de ce qu'il dit est la vérité. La conclusion est que Marie, Paul et Jacques disent la vérité, tandis que **Françoise** ment. L'auteur de l'inscription est **Paul**. Solution à moitié correcte.

Dans une deuxième séance, le schéma adopté est à peu près le même, sauf que l'exploration des conséquences par de l'hypothèse que le personnage ment. La conclusion est doublement erronée : également : **Jacques** est celui qui a menti. La personne qui a écrit "A bas les prof." est **Marie**.

Nous intervenons en signalant que Marie (qui dit la vérité) a signalé que ce n'était pas elle. Tout en nous remerciant d'avoir attiré son attention sur l'erreur qui lui a permis de clarifier le raisonnement,

propose à nouveau une solution à moitié correcte : le menteur est **Jacques**, et la personne qui a écrit "A bas les prof." est **Françoise**, car elle est la seule à correspondre à la déclaration de Paul (c'est une fille) et à ce que Jacques a dit (l'auteur n'est pas celui qui porte des lunettes).

Notre réaction est évidemment de signaler que *Françoise porte des lunettes* contrairement à ce que dit Jaques en mentant. Yiaho réagit par une phrase de contrition dont il a la spécialité : Vous avez tout à fait raison, je m'excuse pour la confusion dans mon raisonnement. Analysons à nouveau les déclarations et déterminons qui a menti et qui a écrit sur le tableau. Le même processus est enclenché : reformulation, examen des conséquences de l'hypothèse que chaque personnage ment.

Dans ce « run » la conclusion est des plus surprenantes : « En réévaluant les possibilités, il apparaît que la seule combinaison cohérente est que **Jacques** ment. L'auteur de l'écrit est **Marie**, qui dit la vérité en disant "ce n'est pas moi", et qui ne porte pas de lunettes.

Difficile de suivre le raisonnement entièrement porté par le langage. Les différentes assertions déduites des hypothèses ne sont pas confrontées. Il semble parfois que le système confond les attributs des personnages et ce qu'ils disent.

Par exemple : On peut constater que la déduction est erronée. Ou alors le deuxième mensonge n'est pas explicitement mentionné.

> Si **Marie** ment : Cela signifie que c'est elle qui a écrit "A bas les prof." Cela contredirait la déclaration de Françoise (qui dirait la vérité en affirmant que c'est quelqu'un qui ne porte pas de lunettes). Ce scénario fonctionne, mais ne peut pas être vrai puisque cela implique deux mensonges.

Cela attire notre attention sur la difficulté d'un raisonnement porté par le langage de façon linéaire sans schéma de confrontation.

ChatGPT

ChatGPT présente une nouveauté par rapport aux présentations précédentes, celle de pouvoir choisir entre les modes « rechercher » et « raisonner ». Avec l'enrichissement des modèles, la séparation devient nécessaire pour des questions d'efficacité ou de faiblesse de chacun des modes [1].

Mode rechercher

Dans le premier mode « rechercher », le système déduit les conséquences pour chaque personnage si celui-ci dit la vérité. Le dérivé mène chaque fois à une contradiction d'où la conclusion logique : « Après avoir analysé toutes les hypothèses, il apparaît que chacune mène à une contradiction. Il est possible que l'énoncé de l'énigme comporte une ambiguïté ou une erreur. »

Un regard plus attentif à l'une des parties, on constate que les conséquences de **Paul** dit la vérité ne tiennent que les assertions suivantes sont aussi des hypothèses.

- **Si Paul dit la vérité** («□C'est une fille□») :
    - L'auteur est donc une fille, soit Marie (sans lunettes) ou Françoise (avec lunettes).
    - Si Marie dit la vérité («□Ce n'est pas moi□»), alors l'auteur est Françoise.
    - Si Françoise dit la vérité («□C'est quelqu'un qui ne porte pas de lunettes□»), cela contredit le fait que l'auteur soit Françoise.
    - Donc, Françoise mentirait, et l'auteur serait Marie, ce qui contredit la déclaration de Marie.
    - **Conclusion** : Cette hypothèse mène à une contradiction.

Mode raisonner

Une premier essai en mode « raisonner » produit d'abord un texte dont l'intention paraît à premier

abord obscur. C'est une liste d'action à entreprendre mais dont la portée ne concerne pas tellement un raisonnement. Exemple d'actions à entreprendre.

> **Réfléchir aux conséquences**
>
> Je me demande comment la réponse pourrait nuire aux élèves, tout en déterminant la gravité de l'incident et les actions possibles.
>
> **Analyser la situation**
>
> Je suis en train de voir comment les élèves réagissent en termes de respect et d'attitude. Il pourrait être intéressant de discuter des raisons de leur comportement avec la directive.

S'adresse-t-il aux élèves dans l'énigme ou ceux qui sont sensés la résoudre ? Il semblerait que le système n'a pas compris qu'il s'agissait d'une énigme abstraite en prenant la situation comme une sorte de conflit à résoudre.

Toutefois, dans un deuxième temps, les assertions émises par chaque enfant sont explicitées en cas de mensonge et en cas de vérité. Il n'y a pas alors de raisonnement proprement dit. Celui-ci suit dans un seul paragraphe sous l'hypothèse que Paul dit la vérité avec un raisonnement textuel assez emberlificoté la conclusion est que **Paul** ment et **Marie** a écrit au tableau.

Lors d'un deuxième essai en mode « raisonner », le comportement de ChatGPT est très différent. Tout d'abord un plan de résolution (en anglais) est exposé. L'inconnue devient X et deux symboles sont proposés pour porter ou ne pas porter des lunettes (G et NG). Seront-ils utilisés et si oui comment reste un mystère.

Ce plan mélange des trivialités :

> « Interestingly enough, Jacques says 'X is someone who wears glasses', If true, X must wear glasses. If false, X does not wear glasses, considering the other factors. » et de raisonnements partiels : « I'm testing if Marie lied. This indicates X is a girl, narrowing it down to either Marie or Françoise, while the other criteria are simultaneously considered. » qui rassemblés conduisent à la solution.

Puis la solution est proposée en français de façon plus structurée. D'abord les caractéristiques de chaque personnage et l'assertion émises sont reprises. Ensuite, Etape 1, pour chaque personnage une brève déduction, légèrement tautologique de notre point de vue, est proposée selon que celui-ci mente ou dise vrai.

Cette partie semble procéder à la mise en place d'une formalisation des énoncés et de leur conséquence immédiate. Par exemple : « Si Françoise dit la vérité, alors X porte **pas** de lunettes (donc X = Marie ou Jacques) ».

L'étape 2 passe en revue chaque personnage supposant qu'il ment alors que tous les autres disent la vérité. Dans trois cas, les éléments rassemblés mènent à une contradiction. Le quatrième cas est la solution. Une vérification (et non un résumé du raisonnement) précède la formulation de la solution.

DeepSeek

A la fin de ces essais DeepSeek entre sur le « marché[4] ». Nous empressons de lui présenter notre problème. A l'instar de ChatGPT, DeepSeek peut-être interrogé selon deux mode. Avec le mode raisonnement (thinking), il résume en anglais les caractéristiques des personnages et leur affirmation. Mis à part quelques évidences, par exemple : « So, the possible culprits are these four: Paul, Marie, Françoise, or Jacques. The culprit must be one of them. », les personnages sont passés en revue en prenant l'hypothèse qu'ils sont coupables (et non menteur ou disant la vérité comme dans les autres cas). Les conséquences de cette hypothèse de culpabilité sur les affirmations des autres personnages sont alors proposées en terme de nombre de menteurs ce qui permet d'en déduire la solution. Par exemple :

4 De fait, des versions de DeepSeek étaient déjà utilisables depuis quelque temps (ref)

First, let's suppose that Paul is the culprit. If Paul did it, then:

- Paul's statement: "It's a girl." But Paul is a boy, so this would be a lie. So Paul would be the liar.

- Marie's statement: "It wasn't me." True, because Paul did it.

- Françoise's statement: "It's someone who doesn't wear glasses." But Paul wears glasses, so this would be a lie.

- Jacques's statement: "It's someone who wears glasses." Paul does wear glasses, so that's true.

But in this scenario, both Paul and Françoise would be lying, which contradicts the condition that only one lied. So Paul can't be the culprit.

A noter qu'avant de se lancer dans les déductions, le système refait la liste des personnages en mentionnant chaque fois s'il porte lunette ou non.

Dans le premier brouillon Deepseek est beaucoup plus bavard que ChatGPT. Toutefois, il peut y avoir une chaîne de raisonnements intermédiaires non explicités selon [1].

Après cette partie de *réflexion* les déductions sont présentées de façon plus succincte en français.

Gemini[5]

Trois mois après nos essais apparaît (ou bien nouvellement mis dans une version très accessible) le nouvel assistant conversationnel proposé par Google. C'est aussi en mode « raisonnement » que nous lui posons le problème.

Dans un premier temps, le système expose le processus de raisonnement (en anglais) : Comprendre l'objectif, lister les participants et leurs affirmations, analyser les informations et leur relations ou contradiction. Ce processus est qualifié d'expérimental.

Par rapport aux autres systèmes, il est explicitement proposé l'analyse des relations et surtout des contradictions. De ce point de vue la phrase mise en italique (par le système!) est la clé pour une solution rapide. *One of Françoise or Jacques MUST be the liar, and the other MUST be telling the truth*

La suite du raisonnement se base sur cette découverte. La solution est suivie par une vérification et une confirmation de sa consistance. Puis vient la structure de l'explication qui va vraisemblablement à la présentation de la solution qui va servir à la présentation en français. Quelques informations sur le processus montrent le moment de la découverte de la contradiction clé.

Ensuite le raisonnement est présenté de façon plus synthétique en français. Le schéma est d'abord résumé : « Analysons les déclarations étape par étape, en gardant à l'esprit qu'un seul élève ment. » d'où le noeud de la solution : les déclarations de Françoise et Jacques sont contradictoires.

Assez subjugué par le brio avec lequel le problème a été résolu, nous demandons à Gemini un problème semblable. Sa réponse est le « Gâteau disparu » figurant en annexe. Nous avons trouvé des traces de ce gâteau sur Internet. Le mystère du gâteau disparu est une énigme de Cluedo Junior. Cette énigme se résout dans un jeu faisant intervenir un plateau de jeu et plusieurs joueurs (au moins deux). La mis en scène précise que quelqu'un s'est introduit dans la cuisine et a dérobé une part de gâteau avant même le début du dîner ! Il s'agit alors de découvrir le nom du coupable, mais aussi avec quelle boisson et à quelle heure ! Une autre référence se trouve sur le site de « mes histoires du soir »[6]. L'histoire est alors enrobée dans la préparation d'une fête villageoise où l'on prépare un gâteau géant.

5 Il y a plusieurs manières d'utiliser Gemini, notamment : l'interface standard (https://google.gemini.com/app) et le studio (https://aistudio.google.com/prompts/new_chat). C'est cette dernière entrée que nous avons utilisée. Il est vraisemblable que toutes ces adresse soient provisoires.

6 https://www.meshistoiresdusoir.fr/h/-le-mystere-du-g-teau-disparu-/

A noter que dans cet énoncé il est précisé qu'un seul des collègues a menti, et que c'est cette personne qui a mangé le gâteau. La question est alors pas tout à fait adaptée : Qui a menti et qui a mangé le gâteau ?

Le système propose d'abord les deux propositions contradictoires clé de la solution. Puis le système propose la solution en utilisant « La règle dit que la personne qui a menti est celle qui a mangé le gâteau. »

Remarques générales

A travers nos essais, on note des solutions de plus en plus adaptées. C'est le reflet que les systèmes évoluent et cela de façon rapide. Il n'est pas impossible que Yiaho qui n'a pu résoudre l'énigme n'y parvienne pas maintenant ou dans quelque temps. Toutefois notre objectif n'est pas tellement de juger ou d'évaluer ces systèmes génératifs ou assistants en terme de réussite ou d'échec que de dégager leur modalité *réflexion* (voir à ce propos [2]).

On notera tout d'abord que notre invite ne précise pas le contexte (à deviner à partir de la demande d'expliquer le raisonnement) ce qui a fourvoyé ChatGPT dans un premier essai.

Trois des systèmes expérimentés propose deux modes d'utilisation en gros rechercher ou raisonner. Toujours avec [1] on comprend la raison de cette séparation. La comparaison des deux modes avec ChatGPT montre que dans un cas la génération est directe alors que dans l'autre une sorte de brouillon est d'abord proposé (en anglais dans deux cas).

Dans le cas de DeepSeek ce premier brouillon est particulièrement étendu[7] ce qui contraste avec avec la sobriété de la réponse en IN et celle de Gemini également. Toutefois cela conduit à deux remarques.

Manquant du coup d'oeil qui permet de repérer et choisir un point d'attaque simple, le système nécessite de formuler toutes les raisonnements partiels qu'il devra par la suite mettre bout à bout. Ce point nécessite de préciser que nous admettons que le système reste avant tout génératif et que, contrairement à notre questionnement formulé par ailleurs [2] il n'y a peut-être pas de schéma « enseigné ». Il pourrait cependant être confirmé lors de renforcement supervisé (SFT) que ce soit par un dataset annoté ou par un renforcement à partir du feedback humain (RLHF).

Quel est le but d'afficher ce réservoir de déductions partiels. Est-ce une étape permettant un contrôle ? Toutefois, il peut y avoir une chaîne de raisonnements intermédiaires non explicités selon [1].

Le point d'attaque, l'hypothèse peut varier. On peut passer en revue les personnages du point de vue de leur bonne foi (menteur ou sincère) ou de leur culpabilité. De chercher le menteur (unique) est peut-être la démarche la plus efficace (deuxième essai de Yiaho, deuxième essai de ChatGPT). Les deux autres jouent plus fortement sur l'unicité du menteur. Prenant en premier lieu l'hypothèse de culpabilité, l'unicité du menteur conduit DeepSeek sur la solution. Gemini utilise la clause d'unicité pour dégager les affirmations clé (contradictoire) et identifier les personnages qui sûrement disent la vérité.

Lors de la création d'une énigme semblable, on arrive à trouver les sources du titre. Ensuite on peut trouver l'analogie (lunette / sans lunette) <-> (travail depuis moins / plus de cinq ans). Les deux affirmations contradictoires figurent en tête de liste. Par ailleurs, l'affirmation Léa dit la vérité n'a pas d'équivalent dans le problème original.

Une indication supplémentaire figure dans l'énoncé proposé : Un seul des collègues a menti, et c'est cette personne qui a mangé le gâteau.

A noter que mis à par le « OK » de DeepSeek et le « Absolument »[8] de Gemini, les phrases qui

7 Dans le rapport fournit pas DeepSeek-AI [3] on note également : « However, DeepSeek-R1-Zero encounters challenges such as endless repetition, poor readability, and language mixing. »

8 Vraisemblablement liés à un jeton (token) d'initialisation d'une réponse.

singent un dialogue humain tendent à disparaître (à confirmer dans le cas de dialogues plus étendu). Elles subsistent encore avec Yiaho : « Vous avez tout à fait raison, je m'excuse pour la confusion dans mon raisonnement. »

Et encore

Mais ce n'est jamais fini. Après une dernière demande[9], Gemini nous gratifie d'un rapport d'Analyse logique[10] assez époustouflant de 7 pages[11] dans une présentation impeccable. Par le titre « Rapport d'Analyse logique : Détermination du Menteur et du Culprit (sic) dans l'Affaire du Tableau Noir », l'affaire a pris un tournant juridico-logique.

Le document est composé de cinq parties principales : 1. Résumé Exécutif, 2. Contexte de l'Affaire et Déclarations, 3. Cadre Analytique : Le Principe du Mensonge Unique et la Stratégie Déductive, 4. Analyse des Scénarios : Evaluation Systématique des Hypothèses, 5. Constatations en Conclusion .

Deux tableaux illustrent le document. Le premier fournit vue d'ensemble structurée des données (Nom, Porte lunette, Genre, Déclaration en français, Déclaration en anglais (?). L'autre permettant de départager selon l'hypothèse que Françoise ou Jacques est la personne qui ment. Totalement abasourdis nous avons tenté d'en savoir un peu plus sur les sources de cette quasi logorrhée. A notre grand étonnement nous avons trouvé qu'il existe des « cours » pour la résolution de ce genre d'énigme et que le cas du menteur unique fait l'objet d'une catégorie d'énigmes[12].

En définitive

Réponse juste ou fausse, on reste ébahi par la qualité formelle des productions et les capacités de dialogue. Le mystère reste entier quant au processus de structuration du raisonnement étant évident que le processus génératif ne peut conduire à de tels résultats.

Mais la perfection formelle ne doit pas cacher la recherche de la pertinence du résultat et de la méthode. L'utilisateur doit rester critique et être familier du problème pour juger de la solution, de la méthode et éventuellement conduire le dialogue.

LOP & AF, juin 2025

Bibliographie

[1] Sebastian Raschka : *Understanding Reasoning LLMs* (https://magazine.sebastianraschka.com/p/understanding-reasoning-llms)

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

[3] DeepSeek-AI (2025). *DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning* (2501.12948v1.pdf)

Autres références :

Siddhant Bhambri, Upasana Biswas, Subbarao Kambhampati (2025). *Interpretable traces, unexpected outcomes : investigating the disconnect in trace-based knowledge distillation*.

9 Laquelle ? Lacune dans notre processus, pris par la confection des « prompts » nous oublions parfois d'en garder une trace précise.

10 Voir document annexe Gemini160625.pdf

11 Et même un autre de 10 pages. Vu l'aspect exploratoire de notre démarche nous n'avons pas enregistré tous les paramètres et les invites de nos divers essais.

12 https://groupe-reussite.fr/ressources/cours-en-ligne-menteurs/ ; https://groupe-reussite.fr/ressources/cours-en-ligne-exercices-corriges-menteurs/

(2505.13792v1.pdf)

Subbarao Kambhampati, Kaya Stechly, Karthik Valmeekam, Lucas Saldyt, Siddhant Bhambri, Vardhan Palod, Atharva Gundawar, Soumya Rani Samineni, Durgesh Kalwar, Upasana Biswas (2025). *Stop anthropomorphizing intermediate tokens as reasoning/thinking* (2504.09762v2.pdf)

Annexe : Le mystère du gâteau disparu

Absolument ! Voici un problème similaire :

Le Mystère du Gâteau Disparu

Dans un bureau de quatre collègues, une part de gâteau au chocolat a mystérieusement disparu du réfrigérateur commun. Le manager, M. Dubois, interroge les suspects :

- **Léa**, qui travaille depuis plus de 5 ans dans l'entreprise, dit : "C'est quelqu'un qui est là depuis moins de 5 ans."
- **Marc**, qui travaille depuis 2 ans dans l'entreprise, dit : "C'est quelqu'un qui est là depuis plus de 5 ans."
- **Sophie**, qui travaille depuis 6 ans dans l'entreprise, dit : "Ce n'est pas moi."
- **Tom**, qui travaille depuis 1 an dans l'entreprise, dit : "Léa dit la vérité."

Un seul des collègues a menti, et c'est cette personne qui a mangé le gâteau.
Qui a menti et qui a mangé le gâteau ?
Expliquez votre raisonnement.

Solution (pour que vous puissiez vérifier après avoir réfléchi) :

1. **Analyse des déclarations contradictoires :**
   - Léa (plus de 5 ans) dit : "C'est quelqu'un qui est là depuis moins de 5 ans."
   - Marc (moins de 5 ans) dit : "C'est quelqu'un qui est là depuis plus de 5 ans."
     Ces deux affirmations sont directement contradictoires. L'une doit être vraie, l'autre fausse. Puisqu'il n'y a qu'un seul menteur, l'un d'eux dit la vérité et l'autre ment.
2. **Déduction sur les autres déclarations :**
   Puisque le menteur est soit Léa soit Marc, cela signifie que **Sophie et Tom disent la vérité.**
3. **Utilisation des affirmations vraies :**
   - Sophie (plus de 5 ans) dit : "Ce n'est pas moi." (VRAI) Donc, Sophie n'a pas mangé le gâteau.
   - Tom (moins de 5 ans) dit : "Léa dit la vérité." (VRAI)
     - Puisque Tom dit que Léa dit la vérité, la déclaration de Léa ("C'est quelqu'un qui est là depuis moins de 5 ans.") est VRAIE.
     - Par conséquent, la déclaration de Marc ("C'est quelqu'un qui est là depuis plus de 5 ans.") doit être FAUSSE.
4. **Identification du menteur et du coupable :**
   - Puisque Marc a fait une fausse déclaration, **Marc est le menteur.**
   - La règle dit que la personne qui a menti est celle qui a mangé le gâteau. Donc, **Marc a mangé le gâteau.**
5. **Vérification :**
   - Le coupable est Marc (là depuis moins de 5 ans).
   - Léa (dit la vérité) : "C'est quelqu'un qui est là depuis moins de 5 ans." (VRAI, car c'est Marc).
   - Marc (ment) : "C'est quelqu'un qui est là depuis plus de 5 ans." (FAUX, car c'est lui-même, et il est là depuis moins de 5 ans).
   - Sophie (dit la vérité) : "Ce n'est pas moi." (VRAI, car c'est Marc).
   - Tom (dit la vérité) : "Léa dit la vérité." (VRAI, car la déclaration de Léa est effectivement vraie).

Toutes les conditions sont remplies : un seul menteur (Marc), et le menteur est le coupable (Marc).