

Towards Human Readable Explanations

LLM-Solve 2026

Kha Luân Feltus¹, Victor Poucet¹, H       Verhaeghe¹

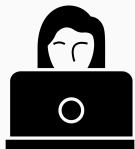
19 July 2026

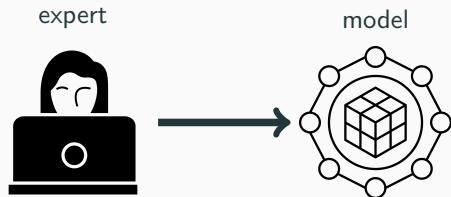
¹ UCLouvain, Belgium

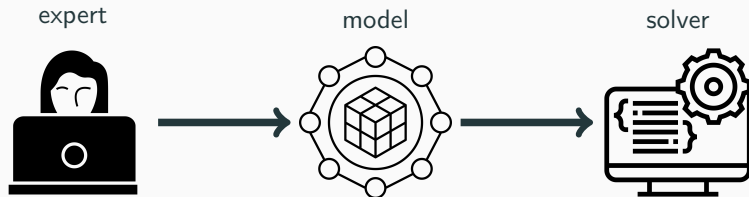
kha.feltus@student.uclouvain.be,
victor.poucet@student.uclouvain.be,
helene.verhaeghe@uclouvain.be

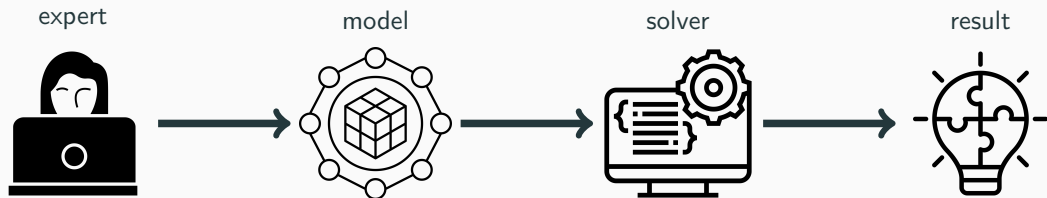


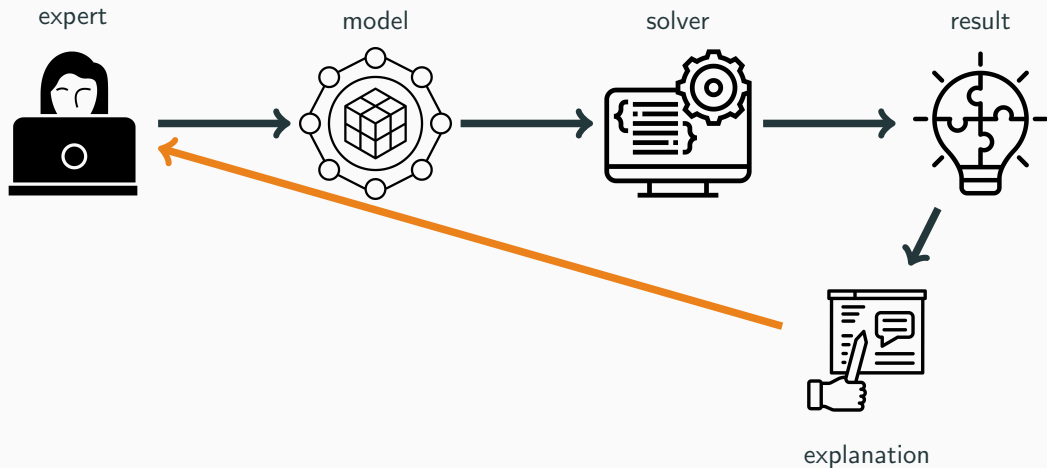
expert

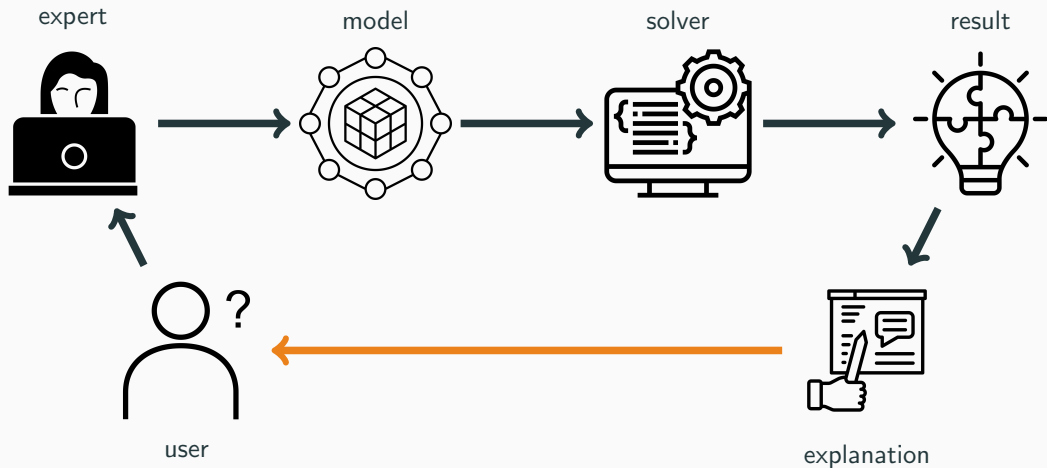


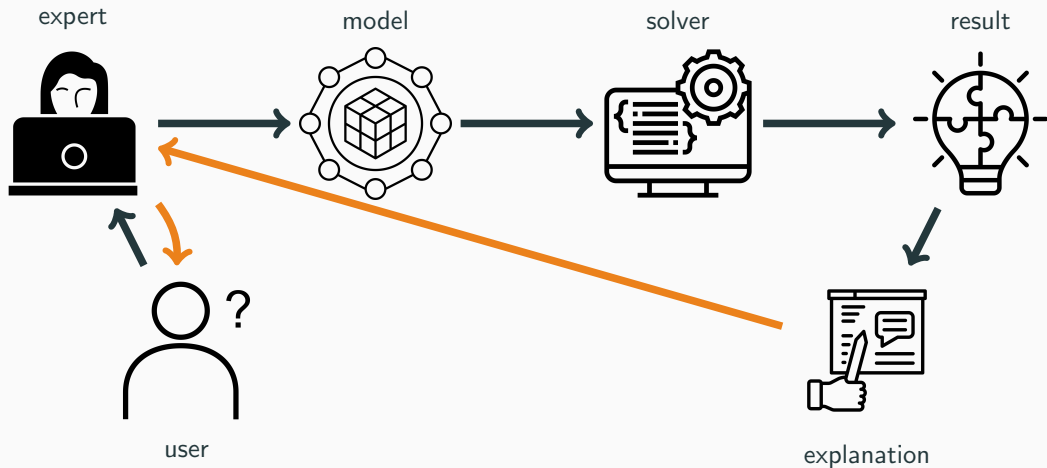


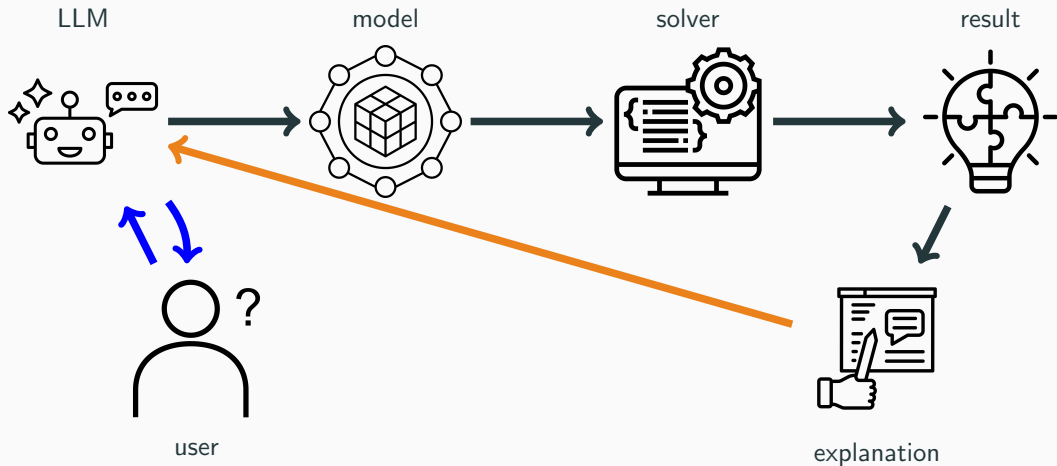


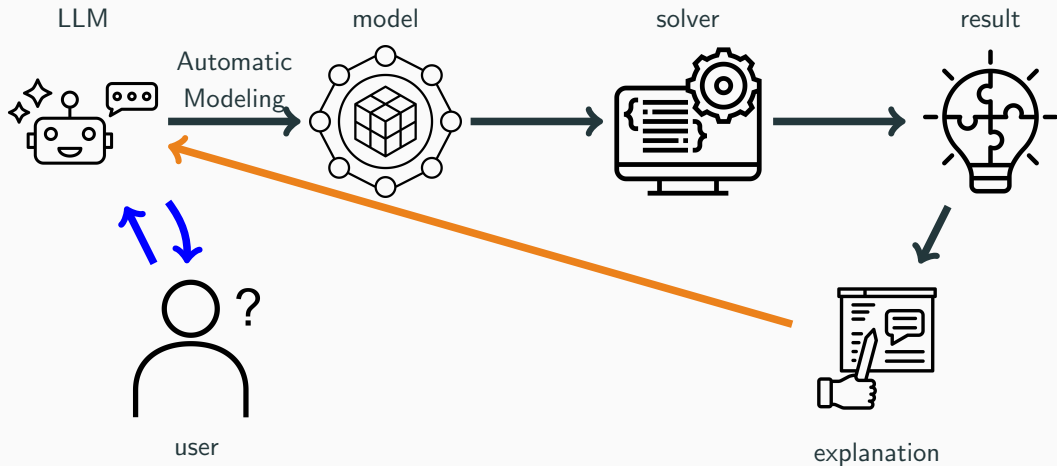


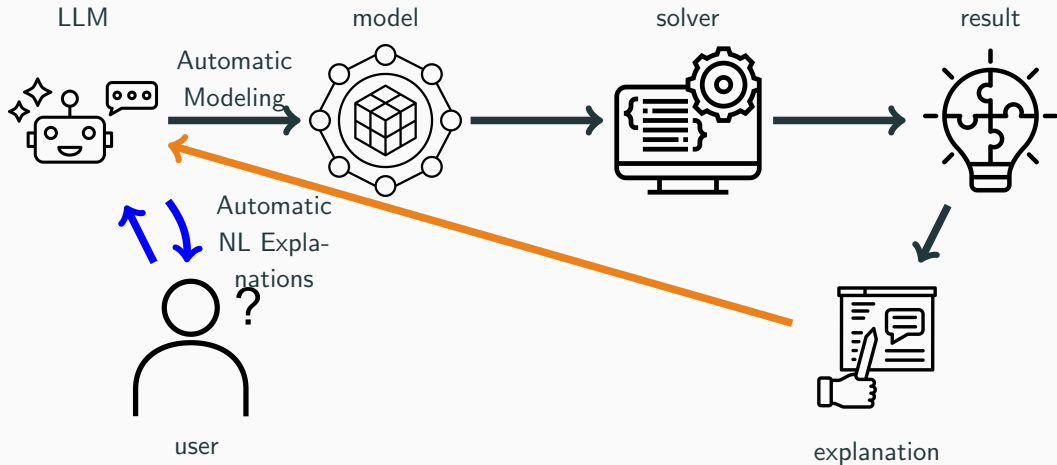












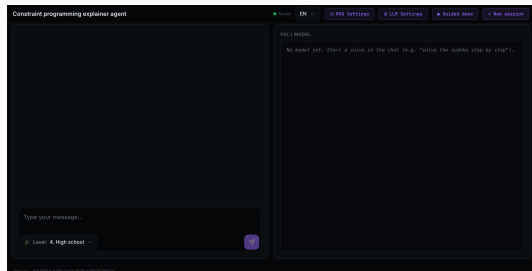
- Explaining the deduction from a decision made in the search tree:
Stepwise Explanations

- Explaining the deduction from a decision made in the search tree:
Stepwise Explanations
- Explaining the dead-end in the search tree:
Minimal Unsatisfiable Subset (MUS)

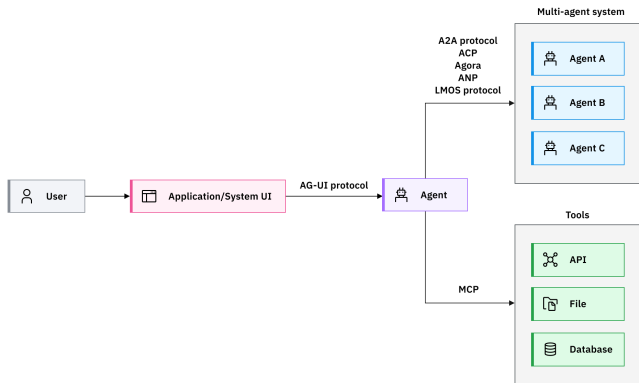
- Explaining the deduction from a decision made in the search tree:
Stepwise Explanations
- Explaining the dead-end in the search tree:
Minimal Unsatisfiable Subset (MUS)

Goal: an interactive ChatBot a non-expert user can chat with to interactively find the solution to their problem and understand the result and process

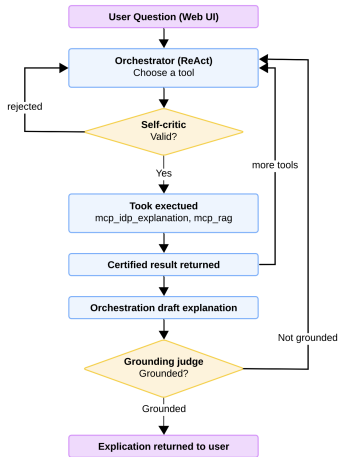
Framework



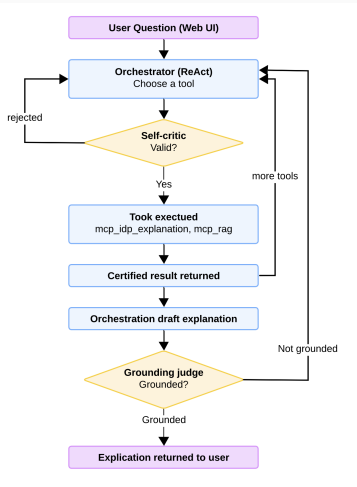
- Three problems pre-loadable:
Sudoku 4x4, Zebra Puzzle, Nurse Rostering
- Stepwise explanation and MUS
- Simulation of search tree



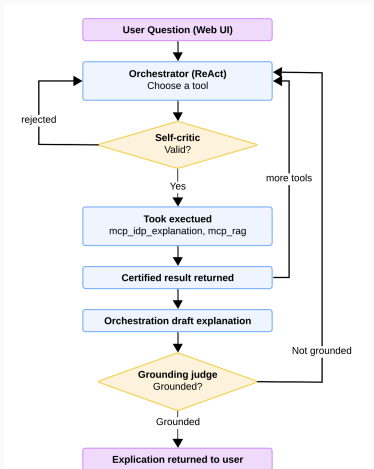
Encapsulation of the symbolic components: MUS computation, stepwise computation, loading problems, getting the states of the variables, selecting a decision,...



- Loop to identify whether a tool is to be called



- Loop to identify whether a tool is to be called
- Auto-evaluation by the LLM to add strenght in the results



Explain that last deduction again, but to a 7-year-old child, in one simple sentence.

```
THOUGHT
🔴 Orchestrator - (tool call)
🟡 Self-Critic - ✓ approved session_state
🔵 Tool Call: session_state - { "session_id":
"0a887559c903468a96c1abd0cdb19d69", "problem_type": "sudoku", "method":
"theory", "known_facts": [ "Value(1, 1) = 1.", "Value(1, 2) = 2.", "Value(1,
3) = 3.", "Value(2, 2) = 4.", "Value(3, 1) = 2..
🔴 Orchestrator - We're playing a game called: 4x4 Sudoku. The game just found
something for sure: 'Value(2, 1) = 3' What we already know in our grid: -
'Value(1, 1) = 1' - 'Value(1, 2) = 2' - 'Value(1, 3) = 3' - 'Value(2, 2) = 4'
- 'Value(3, 1)...
🟡 Judge - ✗ rejected (4/10): The user requested a single, simple sentence
explaining the last deduction to a 7-year-old child, but the proposed response
did not respect this constraint and instead outputted multiple sentences, a
game introduction, bulleted lists, and ...
🔴 Orchestrator - Since our little top-left window already has the toy numbers
1, 2, and 4, the very last empty box in that window has to be the number 3!
🟡 Judge - ✓ accurate (10/10)
```

Since our little top-left window already has the toy numbers 1, 2, and 4, the very last empty box in that window has to be the number 3!

Expertise level

1. Child (age 6-8)
2. Pre-teen (age 10-12)
3. Middle school ✓
4. High school
5. Bachelor's
6. University (beginner)
7. University (advanced)
8. CP / AI expert
9. Formal logic
10. Mathematical proof

- Level 1: *In this row, the numbers 1, 2, and 3 are already there, there is no more room for 4!*

Expertise level

1. Child (age 6-8)
2. Pre-teen (age 10-12)
3. Middle school ✓
4. High school
5. Bachelor's
6. University (beginner)
7. University (advanced)
8. CP / AI expert
9. Formal logic
10. Mathematical proof

Expertise level

1. Child (age 6-8)
2. Pre-teen (age 10-12)
3. Middle school ✓
4. High school
5. Bachelor's
6. University (beginner)
7. University (advanced)
8. CP / AI expert
9. Formal logic
10. Mathematical proof

- Level 1: *In this row, the numbers 1, 2, and 3 are already there, there is no more room for 4!*
- Level 3: *Given the constraint of row, each value can only appear once per row. Row 2 already has 1, 3, and 4, we can thus deduce that cell (2,2) has value 2.*

Expertise level

1. Child (age 6-8)
2. Pre-teen (age 10-12)
3. Middle school ✓
4. High school
5. Bachelor's
6. University (beginner)
7. University (advanced)
8. CP / AI expert
9. Formal logic
10. Mathematical proof

- Level 1: *In this row, the numbers 1, 2, and 3 are already there, there is no more room for 4!*
- Level 3: *Given the constraint of row, each value can only appear once per row. Row 2 already has 1, 3, and 4, we can thus deduce that cell (2,2) has value 2.*
- Level 4: *The column constraint excludes values 1, 3, and 4 as already present. The bloc constraint also excludes 3. The unique remaining value in the domain of (2,2) is thus 2.*

Expertise level

1. Child (age 6-8)
2. Pre-teen (age 10-12)
3. Middle school ✓
4. High school
5. Bachelor's
6. University (beginner)
7. University (advanced)
8. CP / AI expert
9. Formal logic
10. Mathematical proof

- Level 1: *In this row, the numbers 1, 2, and 3 are already there, there is no more room for 4!*
- Level 3: *Given the constraint of row, each value can only appear once per row. Row 2 already has 1, 3, and 4, we can thus deduce that cell (2,2) has value 2.*
- Level 4: *The column constraint excludes values 1, 3, and 4 as already present. The bloc constraint also excludes 3. The unique remaining value in the domain of (2,2) is thus 2.*
- Level 6: *By propagation of the unicity constraints (row, column, bloc), the domain of variable (2,2) is reduced to a singleton $\{2\}$. The assignment $\text{Value}(2,2) = 2$ is a direct consequence of arc consistency.*

can you draw me the sudoku with the known fact in it ?

► THOUGHT

Here is the current state of our 4×4 Sudoku grid!

To help you see our progress clearly, I have marked the **given numbers** (the ones we started with) in normal font, and the **deductions we just figured out together** are **bolded and marked with a star (*)**.



	Col 1	Col 2	Col 3	Col 4
R1	1	2	3	4*
R2	3*	4	.	.
R3	2	.	4*	.
R4	4	.	.	.

Recap of our progress:

- Row 1, Column 4 was filled with **4*** because the other three cells of that row already had 1, 2, and 3.
- Row 2, Column 1 was filled with **3*** because it was the last number missing from the top-left 2×2 block (which already had 1, 2, and 4).
- Row 3, Column 3 was filled with **4*** because it was the only safe spot remaining for a 4 in the third column.*

Framework

Examples

explain me the model

► THOUGHT

Voici une explication simple et imagée de notre modèle. C'est un peu comme la « recette » de cuisine ou la notice de notre jeu de société, écrite de façon à ce qu'un ordinateur puisse la comprendre sans aucune ambiguïté.

Le modèle se divise en trois grandes parties : le **vocabulaire**, les **règles du jeu** (les **contraintes**), et la **forme de la grille**.

1. Le Vocabulaire (Ce que l'on manipule)

Avant de jouer, on définit les objets que l'on va manipuler :

- **Index** : Ce sont les coordonnées de nos cases (lignes et colonnes). Comme notre grille est un carré de 4 sur 4, nous avons les numéros 1, 2, 3, 4 .
- **Val** : Ce sont les chiffres que l'on peut écrire dans les cases, ici également de 1 à 4 .
- **Block** : Ce sont les quatre petits carrés de 2×2 (les blocs). Nous les numérotons de 1 à 4 .
- **Value (La valeur d'une case)** : C'est une fonction mathématique. Si on lui donne une case par sa ligne et sa colonne, elle nous donne le chiffre qui est écrit à l'intérieur (par exemple : la valeur à la ligne 1, colonne 1 est égale à 1).
- **InBlock (L'appartenance à un bloc)** : C'est un outil qui permet de savoir dans quel petit carré de 2×2 se trouve une case.

I want to put $\text{value}(2,3) = 4$

► THOUGHT

J'ai tenté de faire l'hypothèse suivante : placer le chiffre 4 dans la case de la ligne 2, colonne 3 ($\text{Value}(2, 3) = 4$).

Cependant, le solveur nous indique immédiatement que **ce choix est impossible** car il mène à une impasse (le problème devient irrésoluble).

I want to put value(2,3) = 4

► THOUGHT

J'ai tenté de faire l'hypothèse suivante : placer le chiffre 4 dans la case de la ligne 2, colonne 3 (Value(2, 3) = 4).

Cependant, le solveur nous indique immédiatement que **ce choix est impossible** car il mène à une impasse (le problème devient irrésoluble).

Pourquoi est-ce impossible ? (Explication du conflit)

Les seules données qui provoquent ce blocage sont :

- Notre hypothèse : Value(2, 3) = 4
- Une case déjà validée : Value(1, 4) = 4

La seule contrainte enfreinte est :

- La règle du bloc 2×2 : chaque bloc de quatre cases (2 de large sur 2 de haut) doit contenir des chiffres tous différents de 1 à 4.

Si nous plaçons un 4 en ligne 2, colonne 3, nous aurons deux fois le chiffre 4 dans le même bloc situé en haut à droite (qui regroupe les cases de coordonnées (1,3), (1,4), (2,3) et (2,4)).

C'est une contradiction directe avec notre contrainte de bloc !

Retour en arrière :

J'annule donc cette hypothèse et nous revenons en arrière à l'état précédent (retour au nœud de recherche n0) afin de garder notre grille parfaitement juste.*

Continue and reveal the next two forced deductions, one at a time, each with the cells and rule that justify it.

» THOUGHT

Here are the next two forced deductions, presented one at a time with the exact cells and rules used to justify them.

First Forced Deduction

Current problem: Sudoku 4×4

The solver deduced the following fact: $\text{Value}(2, 1) = 3$

The ONLY facts that justify this deduction (minimal cause):

- $\text{Value}(1, 1) = 1$
- $\text{Value}(1, 2) = 2$
- $\text{Value}(2, 2) = 4$

The ONLY constraints involved:

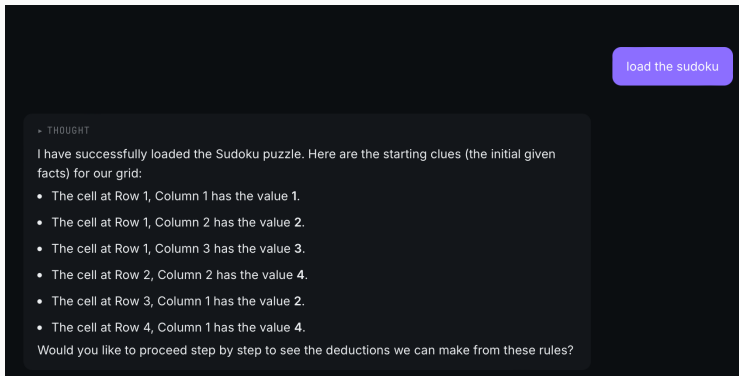
- All four cells in the top-left 2×2 block must contain distinct values from 1 to 4.

Explanation: According to the 2×2 block constraint, the four cells in the top-left block must contain unique numbers from 1 to 4. The facts that row 1, column 1 is 1, row 1, column 2 is 2, and row 2, column 2 is 4 rule out 1, 2, and 4 as options for the final cell in this block. Consequently, the cell at row 2, column 1 is forced to be 3.

Second Forced Deduction

Framework

Offline Demo

A screenshot of a chat interface with a dark background. In the top right corner, there is a purple button with the text 'load the sudoku'. Below it, a message bubble contains the text '» THOUGHT' followed by a paragraph and a bulleted list of clues. The message ends with a question: 'Would you like to proceed step by step to see the deductions we can make from these rules?'.

» THOUGHT

I have successfully loaded the Sudoku puzzle. Here are the starting clues (the initial given facts) for our grid:

- The cell at Row 1, Column 1 has the value 1.
- The cell at Row 1, Column 2 has the value 2.
- The cell at Row 1, Column 3 has the value 3.
- The cell at Row 2, Column 2 has the value 4.
- The cell at Row 3, Column 1 has the value 2.
- The cell at Row 4, Column 1 has the value 4.

Would you like to proceed step by step to see the deductions we can make from these rules?

show a grid form

► THOUGHT

Here is the visual representation of our 4×4 Sudoku grid based on the starting clues:

```
+---+---+---+---+
| 1 | 2 | 3 | . | <- Row 1
+---+---+---+---+
| . | 4 | . | . | <- Row 2
+---+---+---+---+
| 2 | . | . | . | <- Row 3
+---+---+---+---+
| 4 | . | . | . | <- Row 4
+---+---+---+---+
c1  c2  c3  c4
```



(Note: The dots . represent the empty cells that we need to fill using our mathematical deductions.)

Would you like us to find the first forced deduction together using our constraints?

do deductions until you cannot

» THOUGHT

Here are the step-by-step deductions made by the solver up to the point where no more certain steps can be taken:

Deduction 1

Current problem: Sudoku.

The solver deduced the following fact: $\text{Value}(1, 4) = 4$

The ONLY facts that justify this deduction (minimal cause):

- $\text{Value}(1, 1) = 1$
- $\text{Value}(1, 2) = 2$
- $\text{Value}(1, 3) = 3$

The ONLY constraints involved:

- Each row must contain all numbers from 1 to 4 exactly once without repetition.

Explanation: Under the row constraint for Row 1, all numbers from 1 to 4 must appear exactly once. Since the values 1, 2, and 3 are already filled in the first three columns, the fourth column has no other option. Therefore, we must place a 4 in Row 1, Column 4.

Deduction 2

Current problem: Sudoku.

The solver deduced the following fact: $\text{Value}(2, 1) = 3$

The ONLY facts that justify this deduction (minimal cause):

Current Grid State

With these three forced deductions filled in, our grid now looks like this:

+	-	-	+	-	-	+	-	-	+	-	-	+
	1		2		3		4					
+	-	-	+	-	-	+	-	-	+	-	-	+
	3		4		.		.					
+	-	-	+	-	-	+	-	-	+	-	-	+
	2		.		4		.					
+	-	-	+	-	-	+	-	-	+	-	-	+
	4		.		.		.					
+	-	-	+	-	-	+	-	-	+	-	-	+
	c1		c2		c3		c4					



We have now exhausted all direct, certain deductions! To continue solving the puzzle, we must make a **hypothesis** (a trial decision) on one of the empty cells. Which cell would you like to make a guess on, or would you like me to suggest one?

lets try the column 2 row 3 to value 2

► THOUGHT

I made the hypothesis that Row 3, Column 2 is 2. The solver indicates that this state is unsatisfiable because it leads directly to a conflict.

✗ Why is this hypothesis impossible?

- **The facts involved:** We attempted to place 2 at Row 3, Column 2, but Row 1, Column 2 already contains the value 2 as one of our initial starting clues.
- **The constraint involved:** The column constraint states that each column must contain all numbers from 1 to 4 exactly once, meaning no number can be repeated in the same column. Having two cells with the value 2 in Column 2 violates this rule!

Because of this conflict, I am canceling this hypothesis and returning to our previous state at the root node. Row 3, Column 2 definitely cannot be 2!

What would you like to try next instead?

ok, make a decision yourself

► THOUGHT

I have taken the initiative to make a couple of strategic decisions (hypotheses) to help us crack this puzzle!

💡 Hypothesis 1

I made the choice to place 1 at Row 2, Column 3 ($\text{Value}(2, 3) = 1$). The solver indicates that this state is **satisfiable** (it is a valid path!). This choice immediately triggered two forced deductions:

- Row 2, Column 4 is forced to be 2 because Row 2 already had 3, 4, and our newly chosen 1.
- Row 4, Column 3 is forced to be 2 because Column 3 already had 3, 4, and our newly chosen 1.

💡 Hypothesis 2

With those cells filled, I made a second choice to place 1 at Row 3, Column 2 ($\text{Value}(3, 2) = 1$). The solver indicates that this state is also **satisfiable**, and it immediately locks in the remaining numbers to completely solve the grid!

Here is how the final cells fell into place:

1. Row 3, Column 4 is 3 (deduced because Row 3 already contains 2, 1, and 4).
2. Row 4, Column 2 is 3 (deduced because Column 2 already contains 2, 4, and 1).
3. Row 4, Column 4 is 1 (deduced because Column 4 already contains 4, 2, and 3).

Here is how the final cells fell into place:

1. Row 3, Column 4 is 3 (deduced because Row 3 already contains 2, 1, and 4).
2. Row 4, Column 2 is 3 (deduced because Column 2 already contains 2, 4, and 1).
3. Row 4, Column 4 is 1 (deduced because Column 4 already contains 4, 2, and 3).

The Fully Solved Grid!

Applying all these logical deductions, our final completed Sudoku grid is:

```
+---+---+---+---+
| 1 | 2 | 3 | 4 |  <- Row 1
+---+---+---+---+
| 3 | 4 | 1 | 2 |  <- Row 2
+---+---+---+---+
| 2 | 1 | 4 | 3 |  <- Row 3
+---+---+---+---+
| 4 | 3 | 2 | 1 |  <- Row 4
+---+---+---+---+
  c1  c2  c3  c4
```



Every single row, column, and 2x2 block contains the numbers 1, 2, 3, and 4 exactly once!
Outstanding job collaborating on this puzzle! Would you like to try another grid, or do you

can you backtrack?

► THOUGHT

I have backtracked by one level, canceling my second hypothesis ($\text{Value}(3, 2) = 1$) and returning to Node n2 (depth 1), where our active decision is still $\text{Value}(2, 3) = 1$.

show the grid

► THOUGHT

At this backtracked state, we are currently exploring the branch where we made the hypothesis $\text{Value}(2, 3) = 1$.

Here is what our Sudoku grid looks like right now:

```
+---+---+---+---+
| 1 | 2 | 3 | 4 | <- Row 1
+---+---+---+---+
| 3 | 4 | 1 | 2 | <- Row 2
+---+---+---+---+
| 2 | . | 4 | . | <- Row 3
+---+---+---+---+
| 4 | . | 2 | . | <- Row 4
+---+---+---+---+
c1  c2  c3  c4
```



Conclusion

- LLMs are good at explaining math/code to users
- Interactions with the users as a feature
- Automatic adaptiveness to the user knowledge
- Not reliable 100% but mechanisms can improve reliability (e.g., symbolic-based tools through MCP, judges, RAGs)

Thank you for listening!

Any questions?