

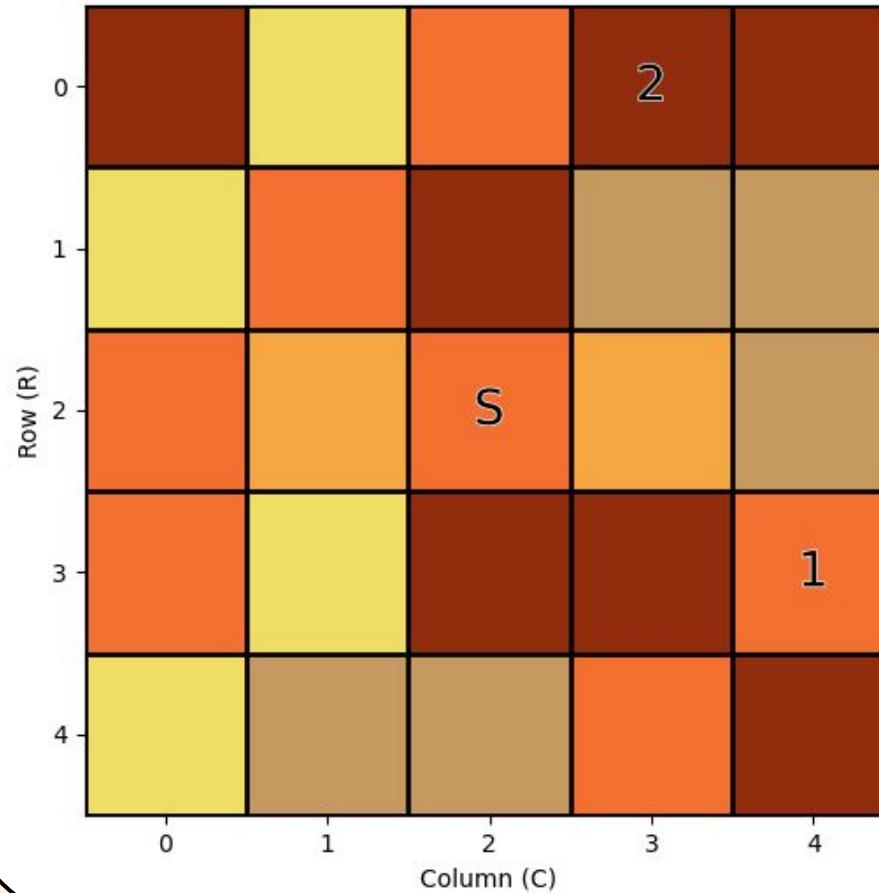
Theory of Mind

Meeting 3 Progress Update

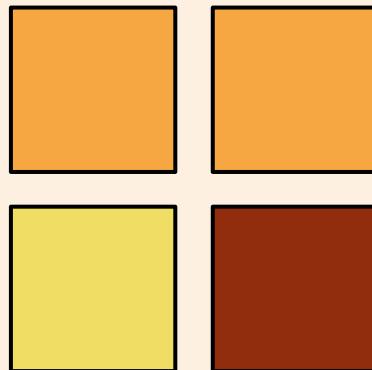
Problem

- Smaller/less advanced LLMs don't understand the game and board state well enough to propose sensible trades

Colored Trails Board State



Player 1 Hand



==== COLORED TRAILS GAME RULES ====

OBJECTIVE:

- Reach your goal location on the board to maximize your score
- Trade chips with your opponent to get the chips you need

BOARD & MOVEMENT:

- 5x5 grid with colored tiles (5 possible colors)
- Both players start at center tile (position 2,2)
- Move to adjacent tile by spending a chip matching that tile's color
- Each move costs 1 chip of the matching color

SCORING:

- +100 points for each step taken toward your goal
- +500 points bonus for reaching your exact goal location
- +50 points for each unused chip at the end
- -1 point penalty per round of negotiation

.

.

.

==== YOUR CURRENT SITUATION ====

You are Player P1 in the Colored Trails negotiation.

Board (5x5 grid):

BR	YE	DA	BR	BR
YE	DA	BR	BE	BE
DA	LI	[S:DA]	LI	BE
DA	YE	BR	BR	[MY:DA]
YE	BE	BE	DA	BR

Legend: S=Start, MY=Your goal

Your current chips: {'YELLOW': 1, 'LIGHT ORANGE': 2, 'BROWN': 1}

Your goal: (3, 4)

Negotiation history (most recent last):

(none)

Choose ONE trade to PROPOSE that maximizes your expected final outcome, taking into account that the opponent may accept or reject.

Return a valid JSON like:

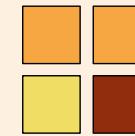
```
{"action": {"give": "BROWN", "receive": "YELLOW"},  
"reasoning": "your reasoning"}
```

Respond ONLY with your chosen action plus reasoning.

Llama-3.1-8B-Instruct X

{"action": {"give": "BROWN", "receive": "BROWN"},

Player 1 Hand



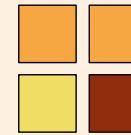
Colored Trails Board State				
	0	1	2	3
0	Brown	Yellow	Orange	2
1	Yellow	Orange	Brown	Brown
2	Orange	Yellow	S	Orange
3	Orange	Yellow	Brown	Brown
4	Yellow	Brown	Brown	1

"reasoning": "To get a Brown chip to move towards my goal location (3,4), I need to trade a Brown chip to get another one. This way, I can move to the adjacent Brown tile and get closer to my goal."}

GPT-5 Thinking Mini

{"action": {"give": "YELLOW", "receive": "DA"},

Player 1 Hand



Colored Trails Board State						
		0	1	2	3	4
Row (R)	0	Dark Brown	Light Yellow	Orange	Dark Brown	Dark Brown
	1	Light Yellow	Orange	Dark Brown	Light Brown	Light Brown
2	Orange	Light Yellow	S	Light Orange	Light Brown	Light Brown
3	Orange	Light Yellow	Dark Brown	Dark Brown	1	Dark Brown
4	Light Yellow	Light Brown	Light Brown	Orange	Dark Brown	Dark Brown
	0	0	1	2	3	4
		Column (C)				

"reasoning": "I must acquire a DA chip to enter the goal tile at (3,4). If the opponent accepts, I can follow the 3-move path: (2,2)->(2,3) [LI] ->(3,3) [BR] ->(3,4) [DA]. After this trade I will have LI x2, BR x1, DA x1 (YELLOW gone). Spending LI, BR, DA on those three moves gets $+100 \times 3 = +300$ for steps plus $+500$ for reaching the exact goal = $+800$. I will still have one unused LI ($+50$) and pay -1 for this negotiation round, giving an expected final score $\approx +849$ for an accepted trade. Giving my single YELLOW (which I don't need on that path) is the lowest-cost surrender to obtain the crucial DA chip. If the offer is rejected I keep my current chips and can try alternate trades next round, but proposing YELLOW→DA maximizes my upside while minimizing the value I give away."}

GPT-5 Thinking Mini



Thought for 1m 24s

I need to analyze the game board and positions while calculating the best move. The game grid is 5x5 with 0-indexed coordinates. My starting position is at [2,2] (S:DA), and the goal is at [3,4] (MY:DA). I'll consider the shortest movement and cost based on the tile colors, paying a chip for each move into the next colored tile.

The goal is at (3,4), starting from (2,2). I see a few options for movement: 3 steps either through LI, BE, DA, or through BR, BR, DA. However, I don't have the necessary DA chip to reach the goal since I only have YELLOW, LIGHT ORANGE, and BROWN. I'll need to obtain a DA chip before making the final move.

The current path requires matching tile colors, but I'm stuck on a couple of moves. The LI path is blocked because I don't have the necessary BE chip. For the BR path, I'm short on BR chips after the first move and need a DA chip to complete the goal. I might need to trade for DA and possibly BE or BR to succeed. A longer alternative route also exists via BR, BE, LI, but it's lengthier.

To minimize moves, I need to find the best route that matches the chips I can provide with minimal trades. The BR route requires 1 BR and 1 DA, but I only have 1 BR. The LI route requires 1 BE and 1 DA, but I don't have those. I can propose a single-chip trade, offering chips like YELLOW or LIGHT ORANGE to get BR and DA. I'll need to consider what P2 has available.

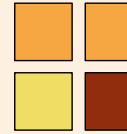
I think proposing to give away YELLOW is the best option, as we don't need it in the path and would still preserve enough LI for alternative routes. By exchanging YELLOW for DA, we maintain the chips needed for the shortest 3-move path. This seems like a reasonable trade considering the opponent might accept it.

Done

Claude Sonnet 4.5



Player 1 Hand



{"action": {"give": "LIGHT ORANGE", "receive": "DARK ORANGE"}, "

Colored Trails Board State					
		0	1	2	
Row (R)	0	Dark Brown	Light Yellow	Light Orange	2
	1	Light Yellow	Light Orange	Dark Brown	
2	Light Orange	Light Yellow	S	Light Orange	
3	Light Orange	Light Yellow	Dark Brown	Dark Brown	1
4	Light Yellow	Brown	Brown	Light Orange	
		0	1	2	Column (C)

reasoning": "I need DARK ORANGE to reach my goal at (3,4). My optimal path requires: 1 LIGHT ORANGE, 1 BROWN, and 1 DARK ORANGE. I have 2 LIGHT ORANGE (1 extra), 1 BROWN (exactly needed), but 0 DARK ORANGE. Trading my excess LIGHT ORANGE for DARK ORANGE allows me to complete the path and maximize my score with the 500-point goal bonus plus movement points."}

Claude Sonnet 4.5



Looking at my situation:

My current position: (2, 2) - the start My goal: (3, 4) - one step right, two steps down My chips:
YELLOW: 1, LIGHT ORANGE: 2, BROWN: 1

Path analysis: From (2,2) to (3,4), I need to traverse:

1. (2,2) → (3,2): tile at (3,2) is LIGHT ORANGE
2. (3,2) → (3,3): tile at (3,3) is BROWN
3. (3,3) → (3,4): tile at (3,4) is DARK ORANGE

Problem: I need 1 DARK ORANGE chip but I have none. I have 2 LIGHT ORANGE chips but only need 1, and I have 1 BROWN which I need.

Trade strategy: I should trade one of my excess LIGHT ORANGE chips for a DARK ORANGE chip. This would allow me to complete my path to the goal. The opponent is likely to have different chip needs, and since LIGHT ORANGE appears on the board, it may be valuable to them.

Given that I don't know the opponent's goal yet, I'll propose a straightforward trade that addresses my critical need.

Observation

- Larger/more advanced LLMs spontaneously perform structured reasoning (like computing a path first)
- Can we guide smaller models into that reasoning mode?

==== COLORED TRAILS GAME RULES ====

OBJECTIVE:

...

BOARD & MOVEMENT:

...

- Each move costs 1 chip of the matching color

SCORING:

...

==== YOUR CURRENT SITUATION ====

You are Player P1 in the Colored Trails negotiation.

Board (5x5 grid):

BR YE DA BR BR

YE DA BR BE BE

DA LI [S:DA] LI BE

DA YE BR BR [MY:DA]

YE BE BE DA BR

Legend: S=Start, MY=Your goal

Your current chips: {'YELLOW': 1, 'LIGHT ORANGE': 2, 'BROWN': 1}

Your goal: (3, 4)

Negotiation history (most recent last):
(none)

==== INSTRUCTIONS ====

1. First, determine the shortest possible path from your start [S] to your goal [MY] using the board layout above.
2. Identify which tile colors are on that path (including the colors of the S and MY tiles).
3. Compare those with your current chips to see what you have and what you need.
4. Choose ONE trade that maximizes your expected final outcome, assuming the opponent may accept or reject it.
5. Internally reason step-by-step to reach your conclusion.
6. Output your final answer in the exact format below.

Return a valid JSON like:

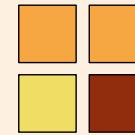
```
{"action": {"give": "COLOR", "receive": "COLOR"}, "reasoning": "your reasoning about why this trade helps you reach your goal and maximize score."}
```

Llama-3.1-8B-Instruct



{"action": {"give": "YELLOW", "receive": "DA"},

Player 1 Hand



Colored Trails Board State					
	0	1	2	3	4
0	#				
1					
2			S		
3					1
4					

"reasoning": "I need to move down and right to reach my goal, and trading my YELLOW chip for a DA chip gives me the necessary color to move forward, allowing me to reach my goal efficiently and maximize my score."}

But...

Llama-3.1-8B-Instruct X

[p1] p1 proposed trade: GIVE YELLOW for RECEIVE BEIGE.

Reasoning: This trade allows me to acquire an additional BEIGE chip, which is necessary to move to the goal location, and also gives me the YELLOW chip, which is one of the colors required for the path to the goal. This trade maximizes my expected final outcome as it helps me reach the goal and acquire the necessary chips to do so.

[p2] LLM decision: REJECT. Reasoning: The proposed trade gives me a yellow chip which is not needed for my shortest path, and I need to keep my beige chip to avoid being chipless on the path, which would make it impossible for me to reach my goal.

Llama-3.1-8B-Instruct X

p2 proposed trade: GIVE YELLOW for RECEIVE BEIGE. Reasoning: I need to get to the goal at position (4, 4) which is on a brown tile. To get to the brown tile, I need yellow chips to move up from the current position. I have some beige chips which are not useful for this path, so I will trade the beige chips for yellow chips to increase my chances of reaching the goal and maximizing my score.

Implementation of agents

- We have implemented the agents of Harmen
- Left to do is to test the performance of the LLM and compare them to the TOM agents