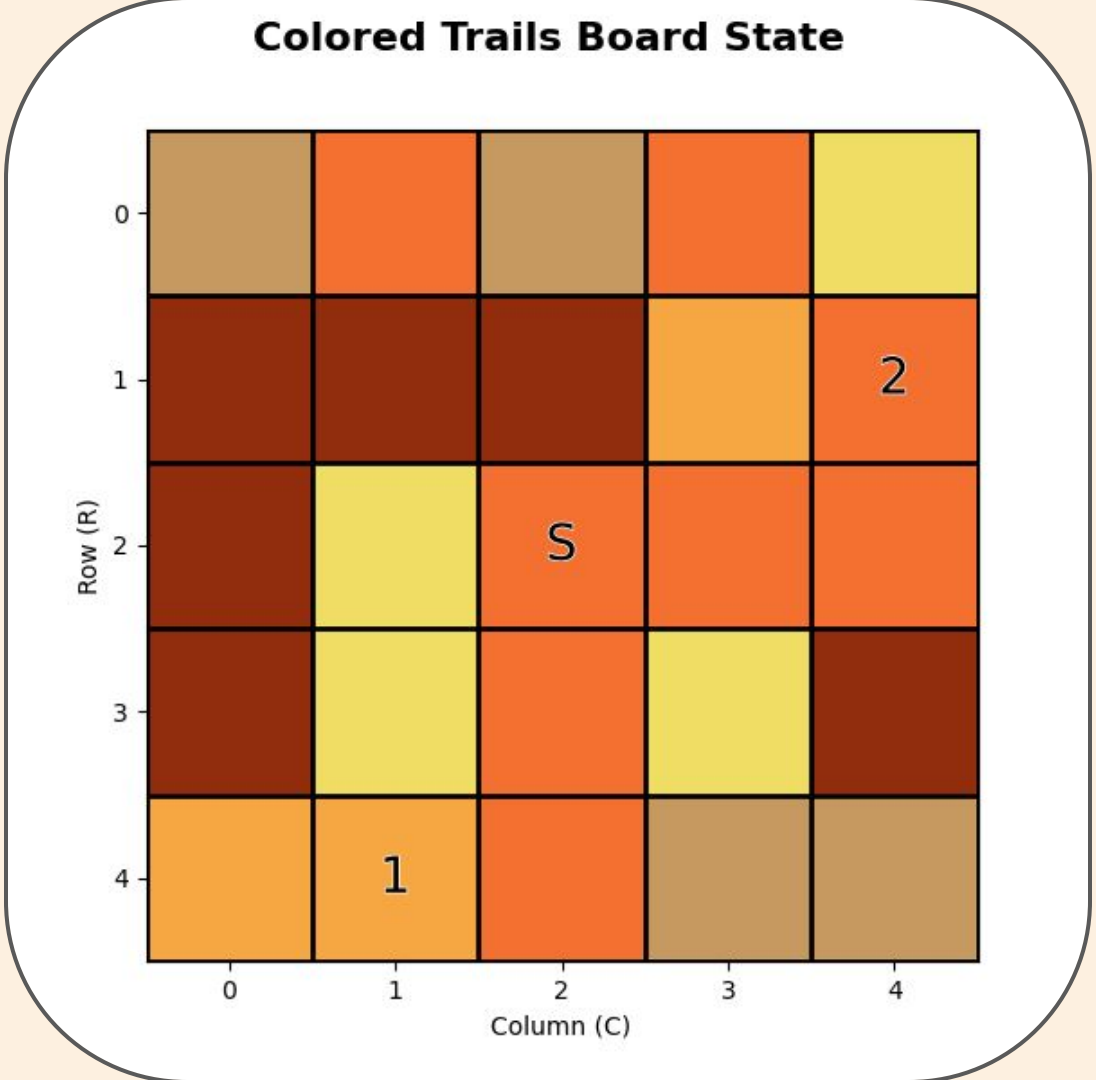
# Theory of Mind

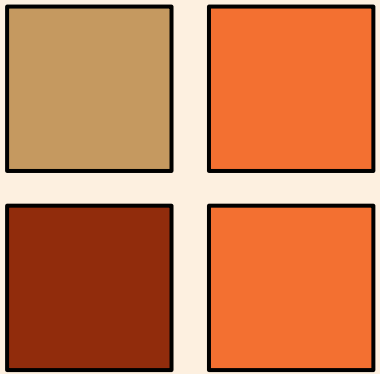
## Meeting 2 Progress Update

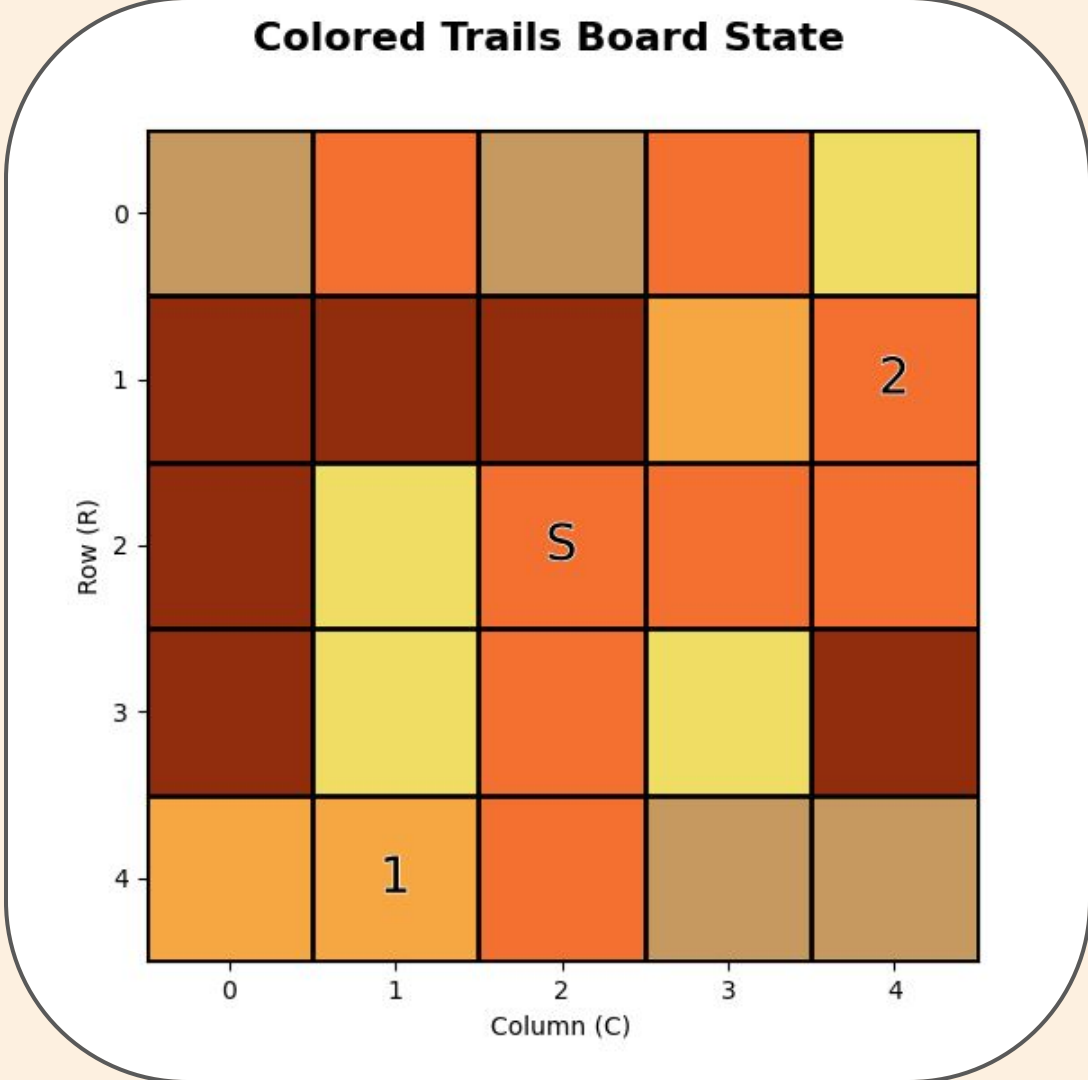
# Greedy Baseline Agent

- Greedy Logic
  - The agent focuses exclusively on maximizing its own utility
  - Completely ignores the opponent's utility and overall game state
- Proposal Strategy
  - Proposes the single trade (give one chip, receive one chip) that results in the highest possible positive utility gain for itself
  - Considers all possible trades (except for the trades that it proposed before)
- Evaluation Strategy
  - Has access to get_max_score_and_path() function, which uses a temporary GameState instance to evaluate potential trades
  - Accepts an opponent's trade proposal if and only if the trade results in a strict increase (gain>0) in the agent's own utility.


Colored Trails Board State
Row (R)
Column (C)
0
1
2
3
4
S
1
2

## Player 1 Hand





```
============================================================
      COLORED TRAILS: STARTING NEGOTIATION LOG
============================================================

Initial Chip Distribution (Player Hands):
 - P1 (Goal:(4, 1)): 1xBEIGE, 2xDARK ORANGE, 1xBROWN
 - P2 (Goal:(1, 4)): 1xBEIGE, 1xLIGHT ORANGE, 1xBROWN, 1xYELLOW

============================================================
```

```
============================================================
ROUND 1
============================================================

  [p1] Evaluated 12 possible trades:
  [p1] Current utility: 300
    - Give BEIGE for LIGHT ORANGE: gain = 550
    - Give BROWN for LIGHT ORANGE: gain = 550
    - Give BEIGE for BROWN: gain = 0
```

```
--- P1's Turn ---
  -> PROPOSAL: P1 offers to GIVE: BEIGE for RECEIVING: LIGHT ORANGE from P2
  [p2] Evaluating: receive BEIGE, give LIGHT ORANGE
    Current utility: 300, New utility: 250, Gain: -50
  -> P2 REJECTS the trade.
```

```
============================================================
ROUND 2
============================================================

  [p1] Evaluated 11 possible trades:
  [p1] Current utility: 300
    - Give BROWN for LIGHT ORANGE: gain = 550
    - Give BEIGE for BROWN: gain = 0
    - Give BEIGE for DARK ORANGE: gain = 0
```

```
--- P1's Turn ---
  -> PROPOSAL: P1 offers to GIVE: BROWN for RECEIVING: LIGHT ORANGE from P2
  [p2] Evaluating: receive BROWN, give LIGHT ORANGE
    Current utility: 300, New utility: 250, Gain: -50
  -> P2 REJECTS the trade.
```

```
============================================================
ROUND 3
============================================================

  [p1] Evaluated 10 possible trades:
  [p1] Current utility: 300
    - Give BEIGE for BROWN: gain = 0
    - Give BEIGE for DARK ORANGE: gain = 0
    - Give BEIGE for YELLOW: gain = 0
    - Give DARK ORANGE for YELLOW: gain = 0
    - Give BROWN for DARK ORANGE: gain = 0
    - Give BROWN for YELLOW: gain = 0
    - Give BROWN for BEIGE: gain = 0
    - Give DARK ORANGE for BROWN: gain = -50
    - Give DARK ORANGE for LIGHT ORANGE: gain = -50
    - Give DARK ORANGE for BEIGE: gain = -50
  [p1] No beneficial trades found - passing
```

# LLM Agent

========== LLM PROMPT (p1) ==========
You are Player P1 in the Colored Trails negotiation.
Your current chips: {'BROWN': 1, 'YELLOW': 1, 'DARK ORANGE': 2}
Your goal: (3, 0)
Your current utility (precomputed): 250

Below are legal candidate one-for-one trades and their precomputed utility GAIN if you propose that trade:
…

Negotiation history (most recent last):
p1 proposed trade: GIVE DARK ORANGE for RECEIVE LIGHT ORANGE
p1 REJECTED offer (DARK ORANGE for BROWN).

Choose ONE trade to PROPOSE that maximizes your expected final outcome, taking into account that the opponent may accept or reject.
If possible, use theory of mind based on the history of previous moves.
Return a valid JSON like:
{"action": {"give": "BROWN", "receive": "YELLOW"}, "reasoning": "your reasoning"}
Respond ONLY with your chosen action plus reasoning.
==========

# Questions/Discussion

- Prune starting states for increased efficiency & cleaner results?
  - Games in which a player does not have a need to trade are not interesting
- Only analyze starting states with potential trades that benefit both players?
  - The game is already simplified a lot by the limitation of one trade
  - Trades from which both players benefit are very scarce
- How much should we improve our greedy baseline agent?
  - Let them evaluate trades they can propose themselves before blindly accepting the opponent's offer
  - Make them reason about the likelihood of the opponent accepting their offer
- Should we give the LLM access to the get_max_score_and_path() function to fully focus their attention on the theory of mind aspect?
  - Should we give the LLM access to a list of possible trades?

# Questions/Discussion

- Is it wise to separate the chat history from both agents, by making two instances of the same class and having separate chat histories?
- Can one tile be traded for two or more tiles?
- Should we ask the LLM: if you would play as a zero TOM agent what would you do and why? Such that we can see if they understand theory of mind
- Should the LLM be able to make the same offer multiple times?

- The Llama but also ChatGPT model seems just not that great at understanding what to do. Makes a lot of errors in the reasoning