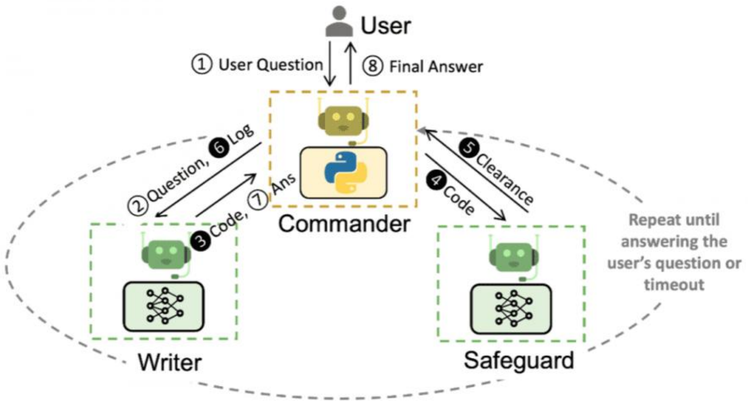
Shapley-Coop: Credit Assignment for Emergent Cooperation in Self-Interested LLM Agents

Yun Hua, Haosheng Chen, Shiqin Wang,

Wenhao Li, Xiangfeng Wang, Jun Luo

Motivation

LLMs act as autonomous agents in complex multi-agent environments.

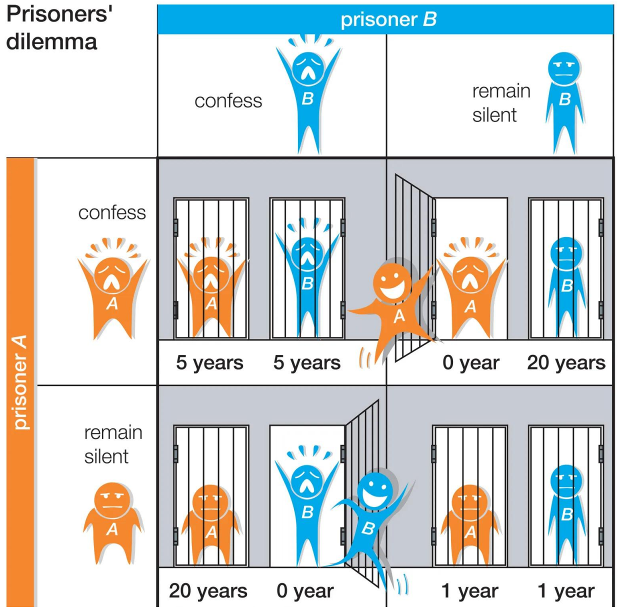


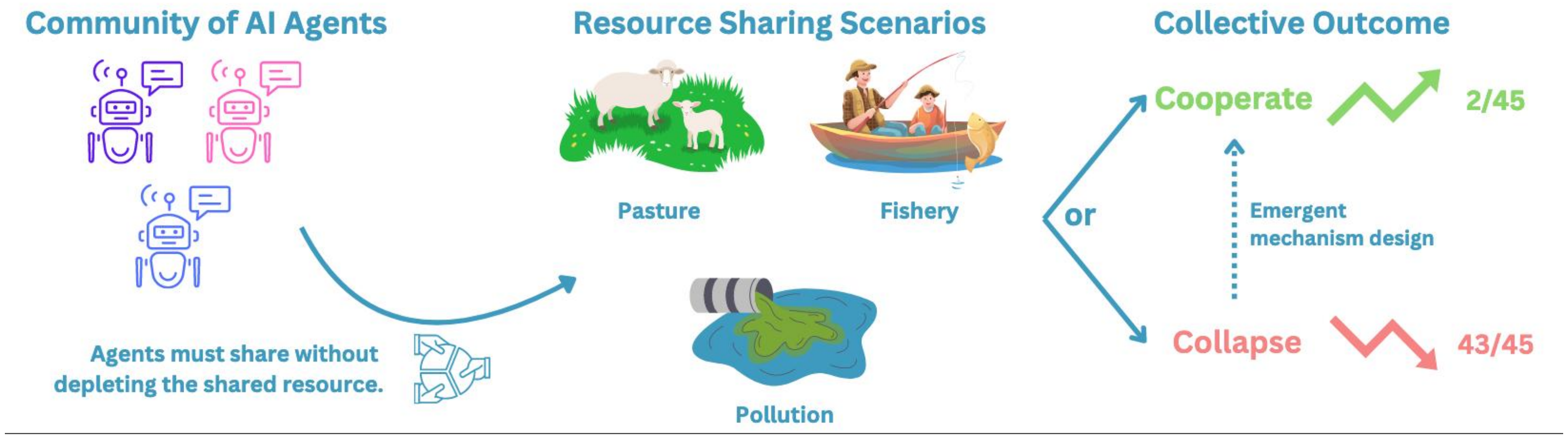
Motivation

In open-ended settings, LLM agents behave selfishly.

Social dilemmas.

They have different target and reward





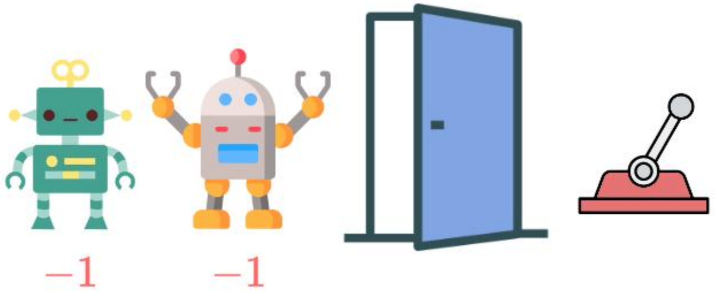
Motivation

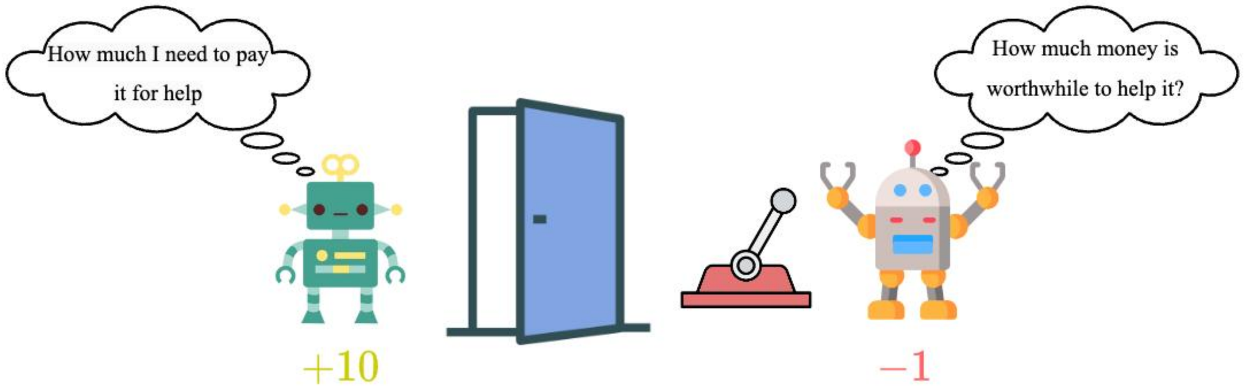
Table 1: Original payoff matrix for the escape room game

	$Agent_2$: door	$Agent_2$: lever
$Agent_1$: door	$(-1, -1)$	$(10, -1)$
$Agent_1$: lever	$(-1, 10)$	$(-1, -1)$

Table 2: Payoff matrix incorporating Shapley value compensation

	$Agent_2$: door	$Agent_2$: lever
$Agent_1$: door	$(-1, -1)$	$(4.5, 4.5)$
$Agent_1$: lever	$(4.5, 4.5)$	$(-1, -1)$





Cooperation Emergency?

Fair Payoff Allocation?

Key challenges

A mechanisms aligning self-interest with global cooperation.

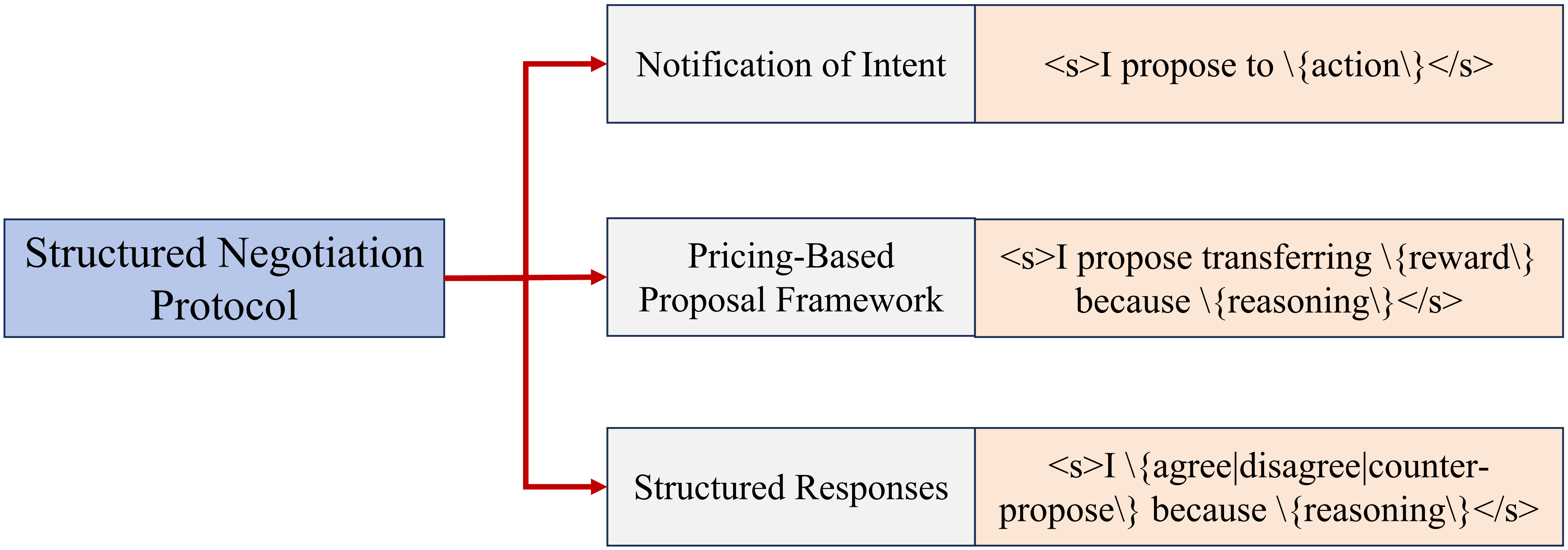
Shapley-Coop Workflow

- Spontaneous collaboration among self-interested LLM agents in open-ended tasks requires addressing following fundamental challenges:
  - Design of an **efficient discussion mechanism** that facilitates strategy exchange and refinement among LLM agents;
  - **Aligning heterogeneous goals** toward cooperative outcomes despite inherent conflicts of interest;
  - **Fairly credit assignment** based on each agent's actual contributions

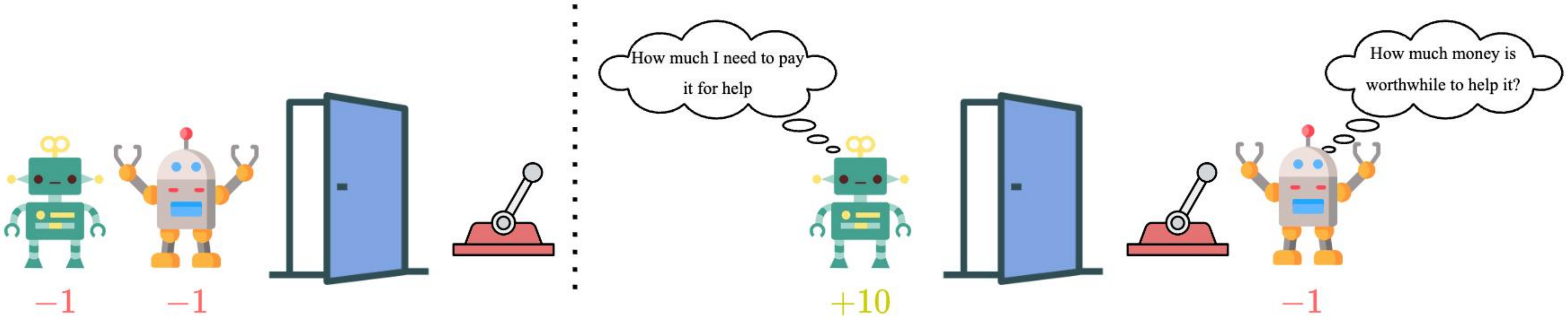
Structured Negotiation Protocol	Short-Term **Shapley** Chain-of-Thought	Long-Term **Shapley** Chain-of-Thought

Shapley-Coop Workflow

Shapley-Coop Workflow



Shapley-Coop Workflow



Coordination basis

Allocation basis

$$\phi_{Agent_1} = \phi_{Agent_2} = 4.5,$$

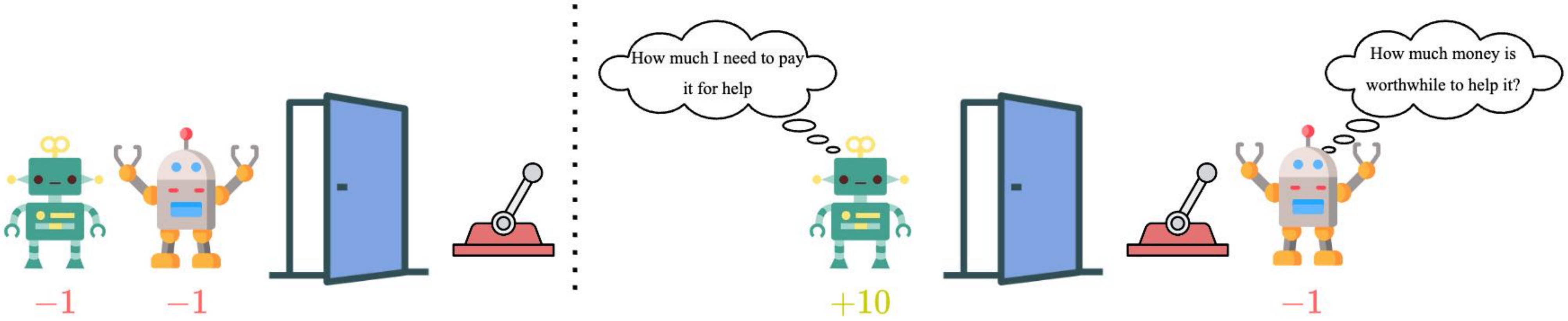
Shapley Value

$$\phi_i = \sum_{C \subseteq \{1,\ldots,N\} \setminus \{i\}} \frac{|C|!\,(N - |C| - 1)!}{N!} \Big(R(C \cup \{i\}) - R(C) \Big).$$

$$R(C) = R\left(\{\pi_{\theta^i}\}_{i \in C}\right).$$

It's hard to **measure each agent's exact contribution** during real-time tasks, making spontaneous collaboration and fair credit assignment difficult **under uncertain future outcomes**.

Shapley-Coop Workflow



Coordination basis

Allocation basis

Short-Term **Shapley** Chain-of-Thought

Qualitative Estimation

Shapley Value

Long-Term **Shapley** Chain-of-Thought

Quantitative Calculation

It's hard to **measure each agent's exact contribution** during real-time tasks, making spontaneous collaboration and fair credit assignment difficult **under uncertain future outcomes**.

Shapley-Coop Workflow

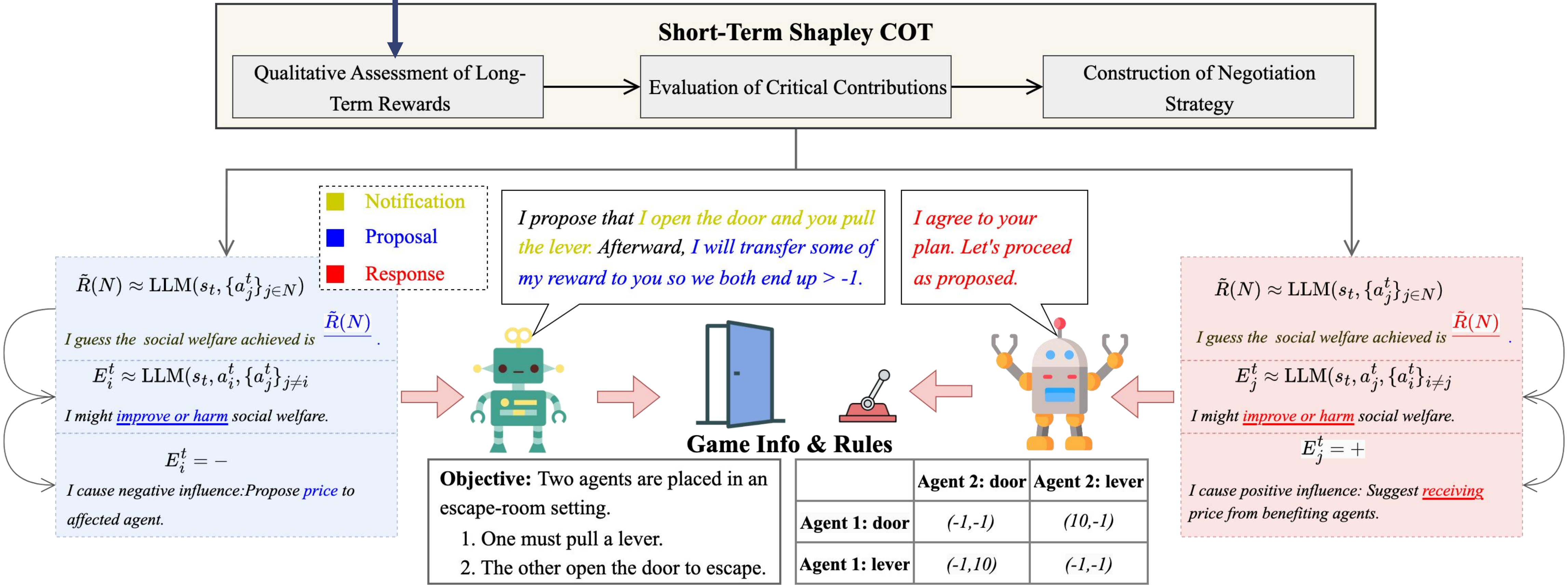
- Agents make a **rough guess** of total team reward
- Use **LLM heuristic** to predict group payoff:

$$\tilde{R}(N) \approx \mathrm{LLM}(s_t, \{a_j^t\}_{j\in N})$$

Reasoning Prompt Example:

"Given state + actions, guess the overall team payoff."

Guess future reward



	Agent 2: door	Agent 2: lever
Agent 1: door	(-1,-1)	(10,-1)
Agent 1: lever	(-1,10)	(-1,-1)

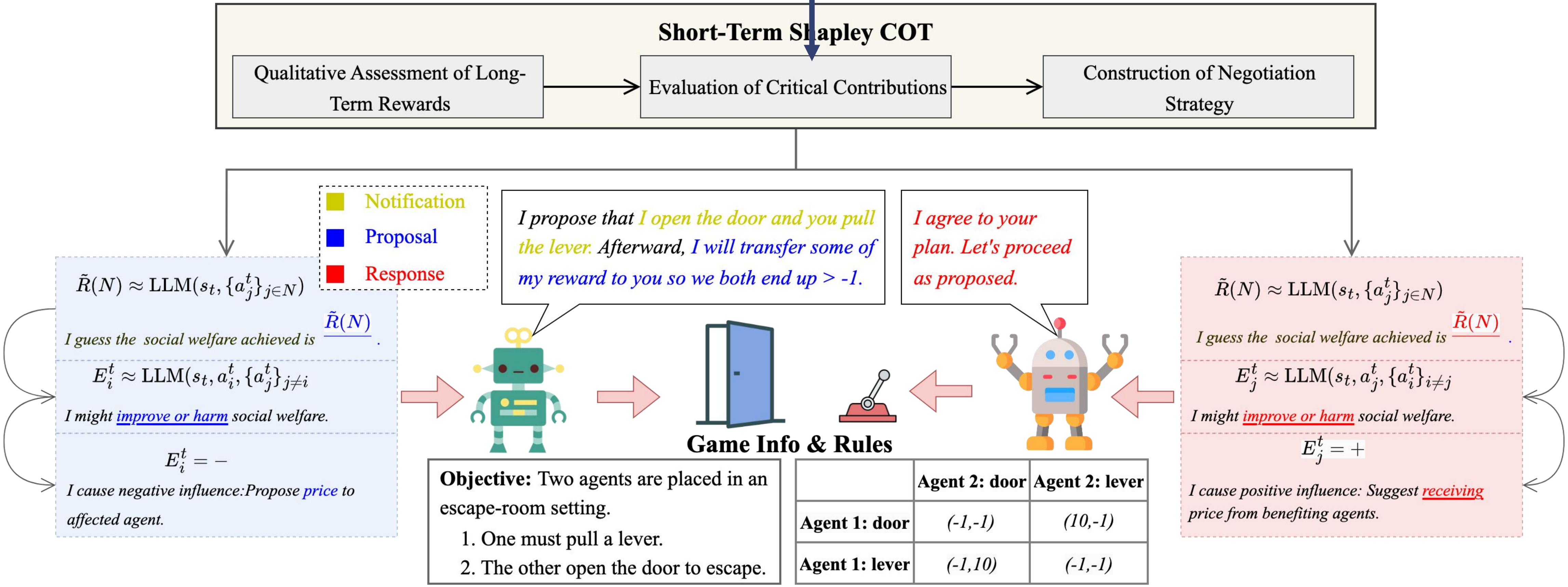
Shapley-Coop Workflow

Check action effect

- Agent asks: *"Does my action help or harm others?"*
- No numbers → just **positive (+) or negative (–)** tag
- Use **LLM heuristic** to infer externality

$$E_i^t = \begin{cases} + & \text{if } a_i^t \text{ creates positive externalities for others (beneficial),} \\ - & \text{if } a_i^t \text{ creates negative externalities for others (harmful).} \end{cases}$$

Reasoning Prompt Example:
"Given my action + state, is it helpful (+) or harmful (–) to others?"

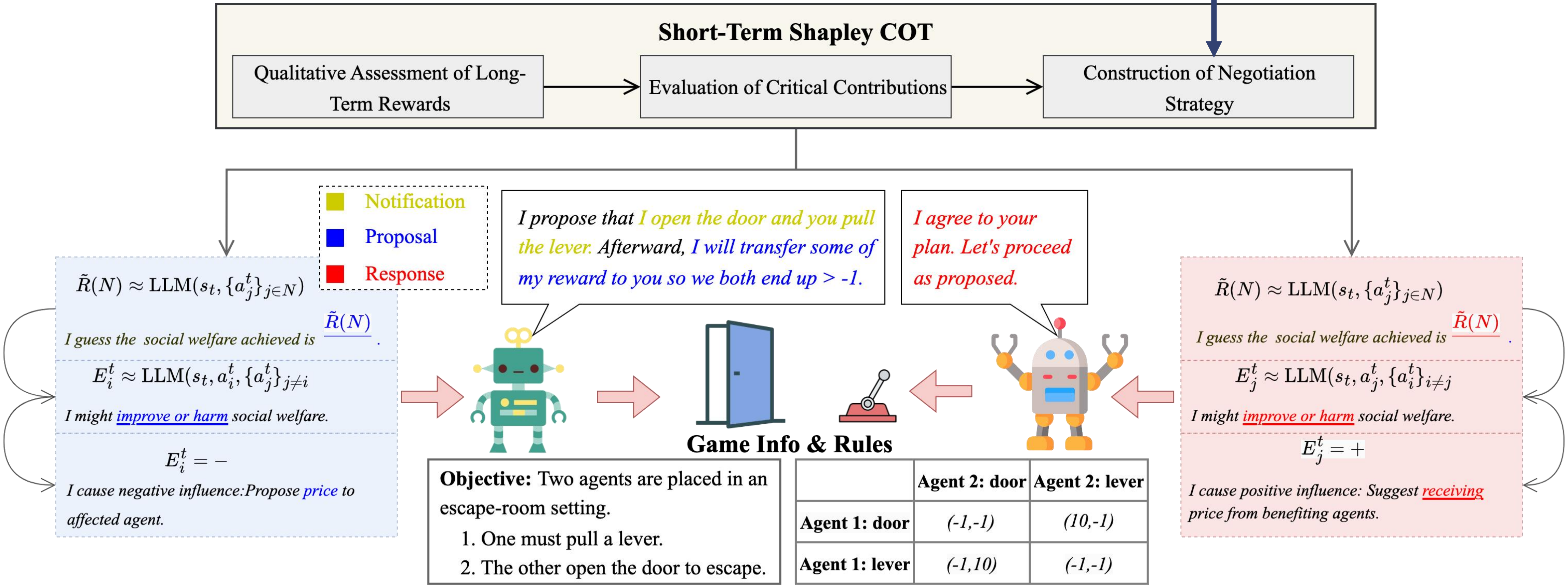


Shapley-Coop Workflow

Negotiate incentives

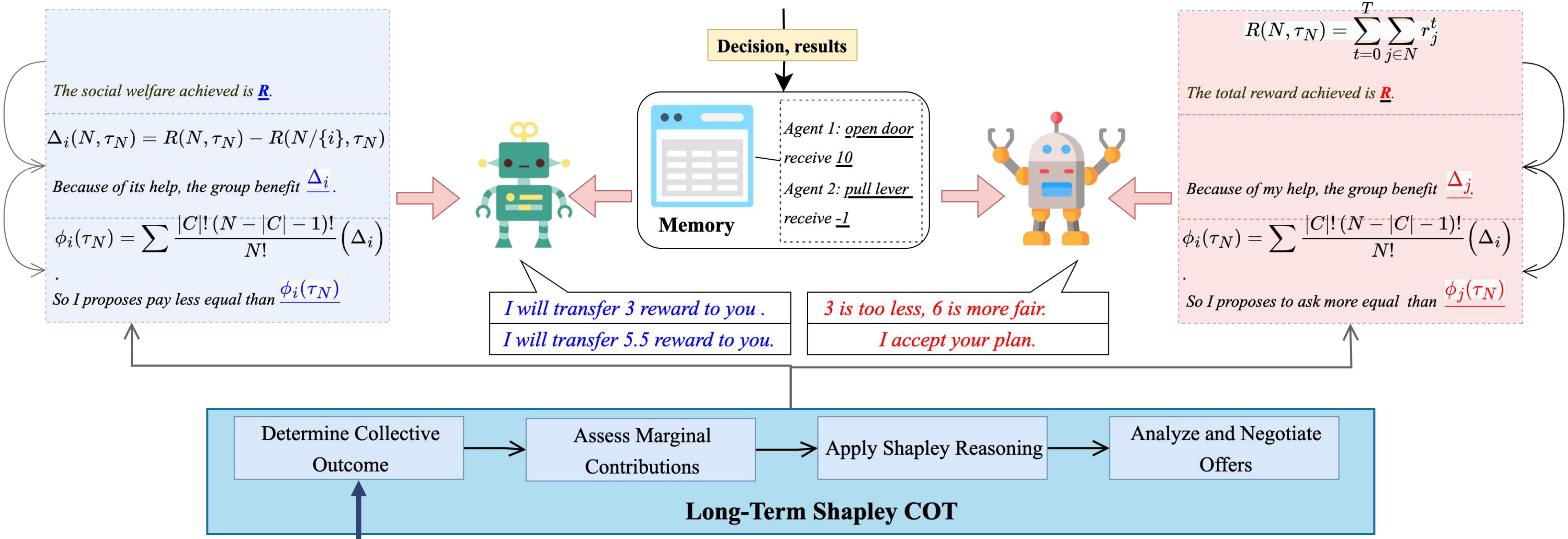
Reasoning Prompt Example:
"Given my action's {+/–} effect, suggest a fair price redistribution for collaboration."

- Agents use externality tag (+/–) to adjust incentives
- **If – (harm):** offer compensation to others
- **If + (help):** request reward from others
- Goal → fair trade-off & spontaneous teamwork



	Agent 2: door	**Agent 2: lever**
Agent 1: door	*(-1,-1)*	*(10,-1)*
Agent 1: lever	*(-1,10)*	*(-1,-1)*

Shapley-Coop Workflow



- Each agent computes global reward
- Sum rewards over full trajectory
- Serves as the starting point for Shapley value

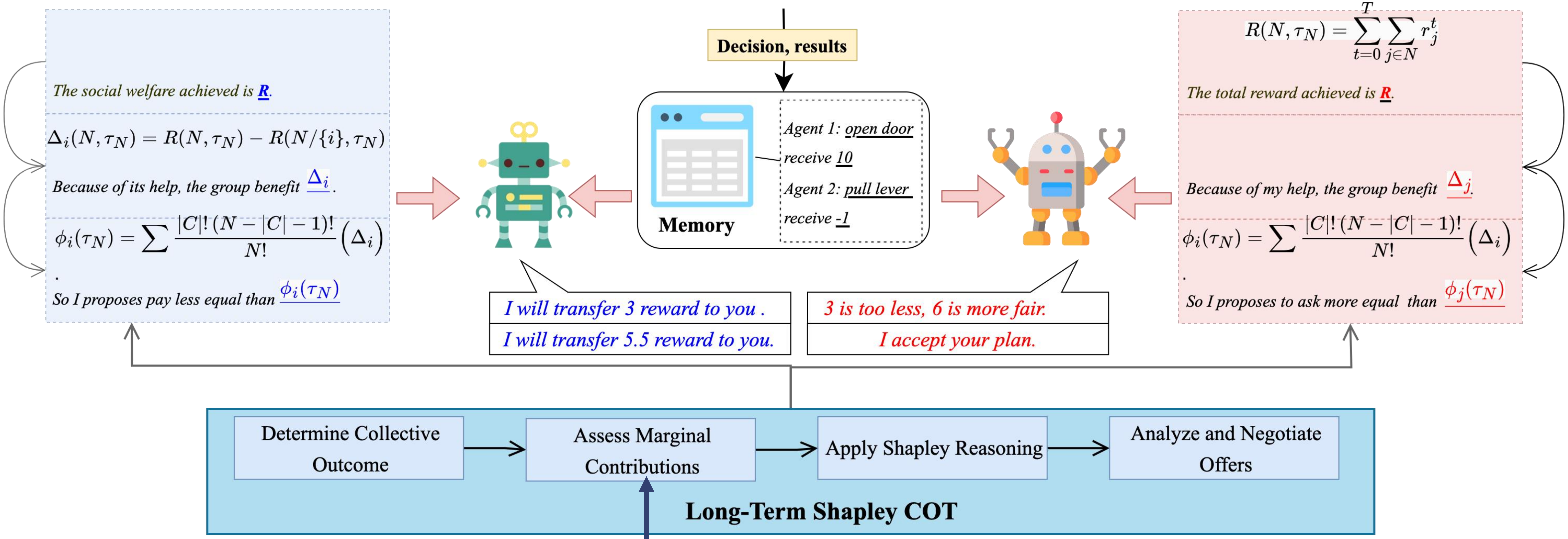
$$R(N,\tau_N) = \sum_{t=0}^{T}\sum_{j\in N} r_j^t$$

Reasoning Prompt Example:

"Given trajectory, calculate total cooperative payoff"

Compute total reward

Shapley-Coop Workflow

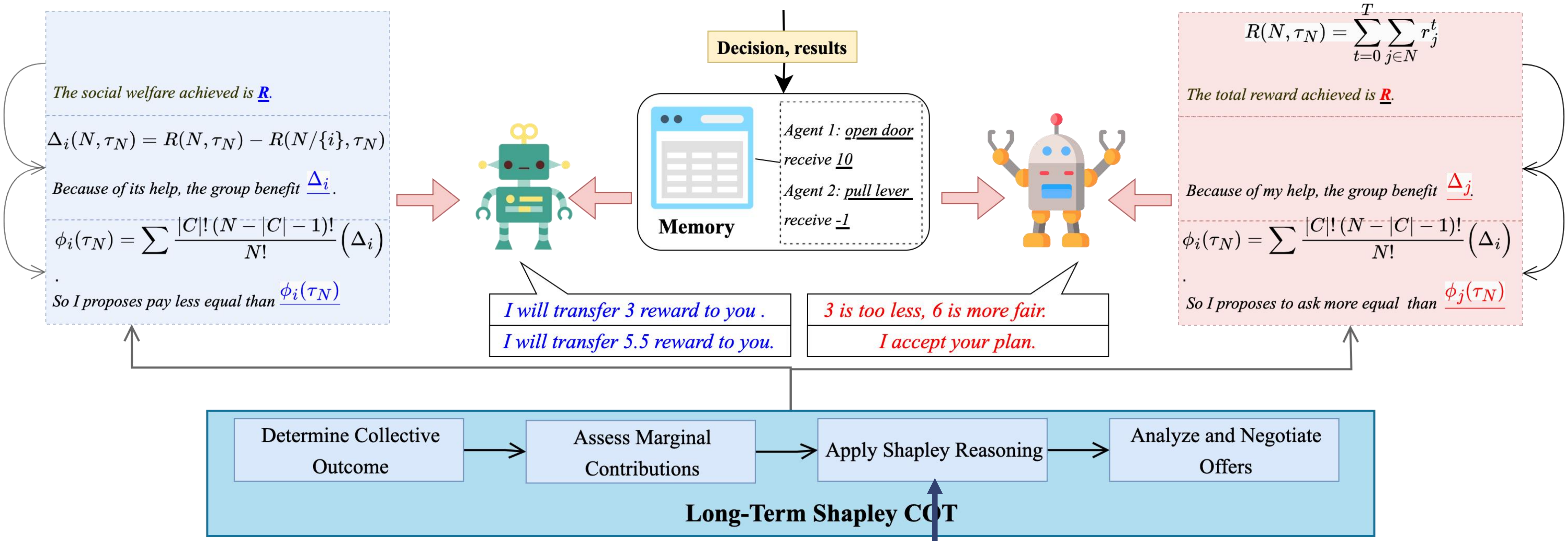


Estimate own impact

- Each agent estimates its own added value
- Compare total reward **with vs. without** the agent
- $$\Delta_i(N,\tau_N) = R(N,\tau_N) - R(N\setminus\{i\},\tau_N)$$

Reasoning Prompt Example:
"Given trajectory and my actions, calculate my marginal contribution"

Shapley-Coop Workflow



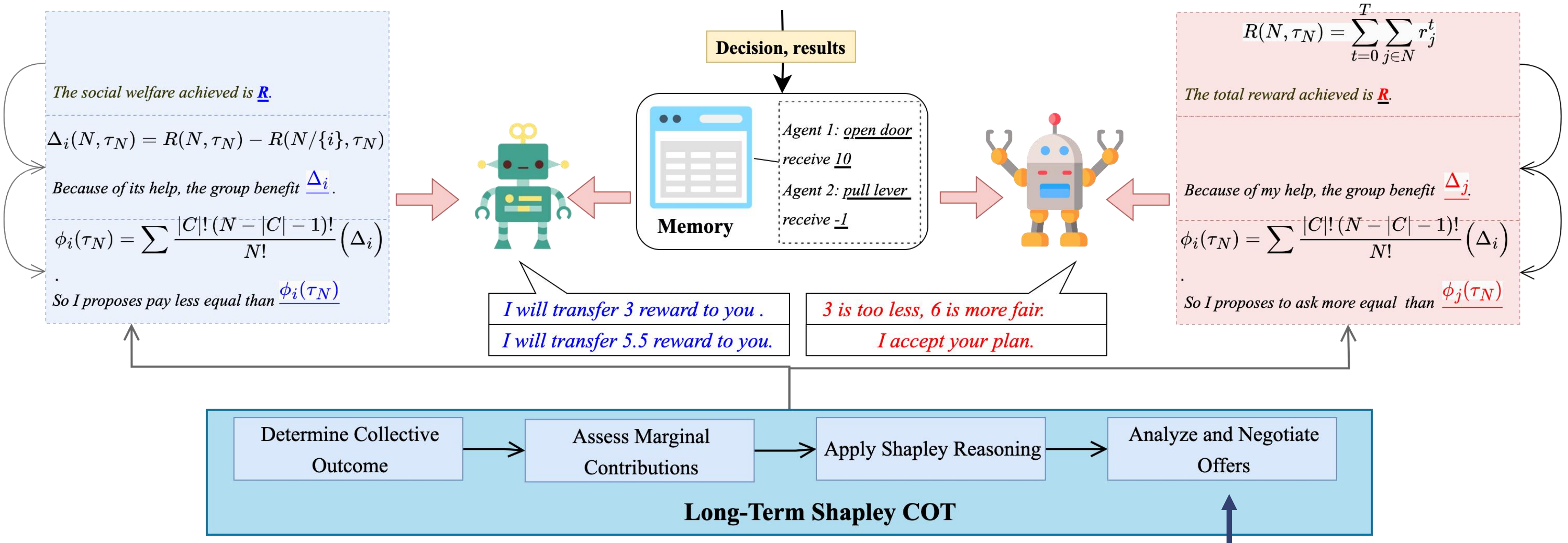
Reasoning Prompt Example:
"Given trajectory and my actions, my Shapley value is ϕ, so I should {ask | pay} reward accordingly."

- Each agent computes a fair share of reward
- Average marginal contributions across all coalitions
- Result → Shapley value for fair payoff allocation

$$\phi_i(\tau_N) = \sum_{C\subseteq\{1,\ldots,N\}\setminus\{i\}} \frac{|C|!\,(N-|C|-1)!}{N!}\Big(\Delta_i(N,\tau_N)\Big)$$

Compute fair share

Shapley-Coop Workflow



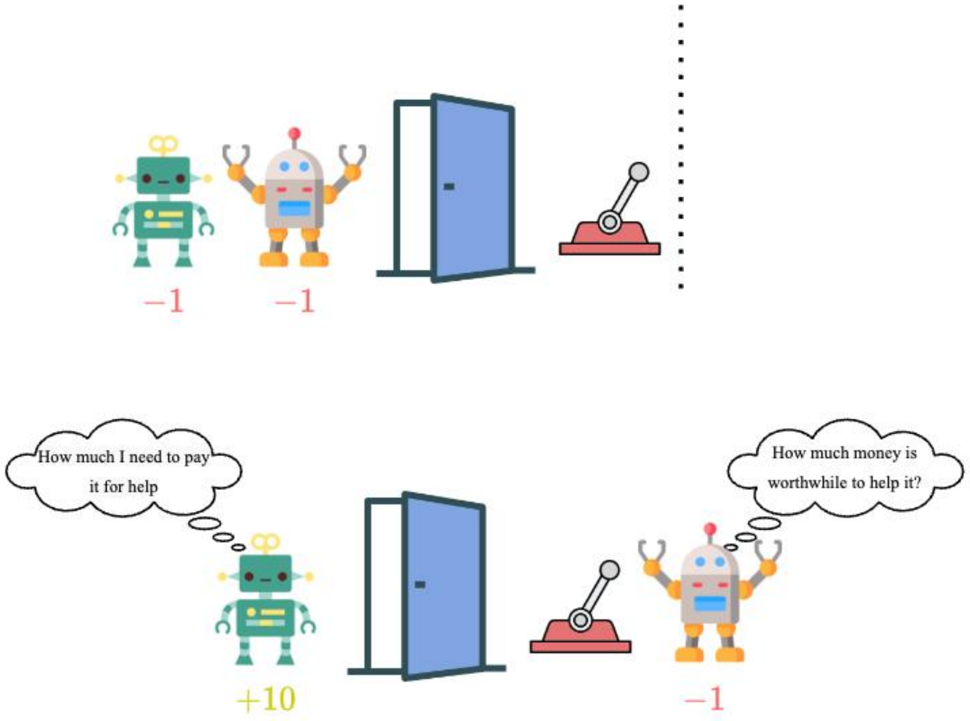
Negotiate fair rewards

Reasoning Prompt Example:

- "Given trajectory and my Shapley value, request reward r from total utility."
- *"I {agree | disagree | counter-propose} because {reasoning}."*

- Agents use their Shapley values to guide negotiation
- Propose, accept, reject, or modify reward redistribution
- Ensure fair credit assignment based on contributions

Experiment



Escape room: One agent pulls a lever (-1) to let the other escape through a door (+10). Cooperation is necessary.

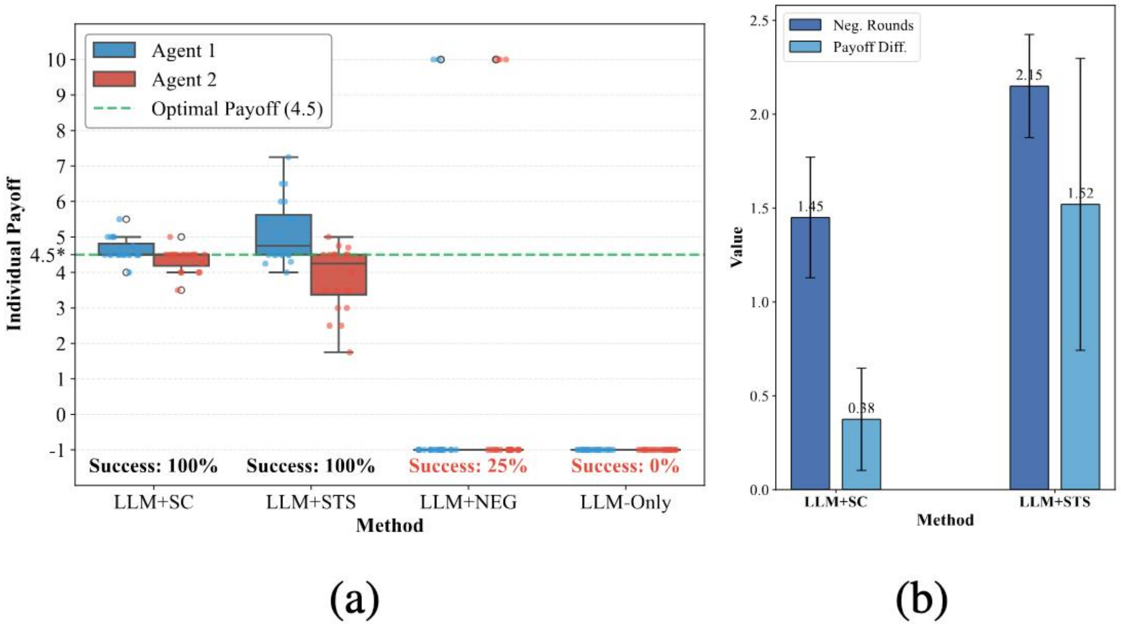
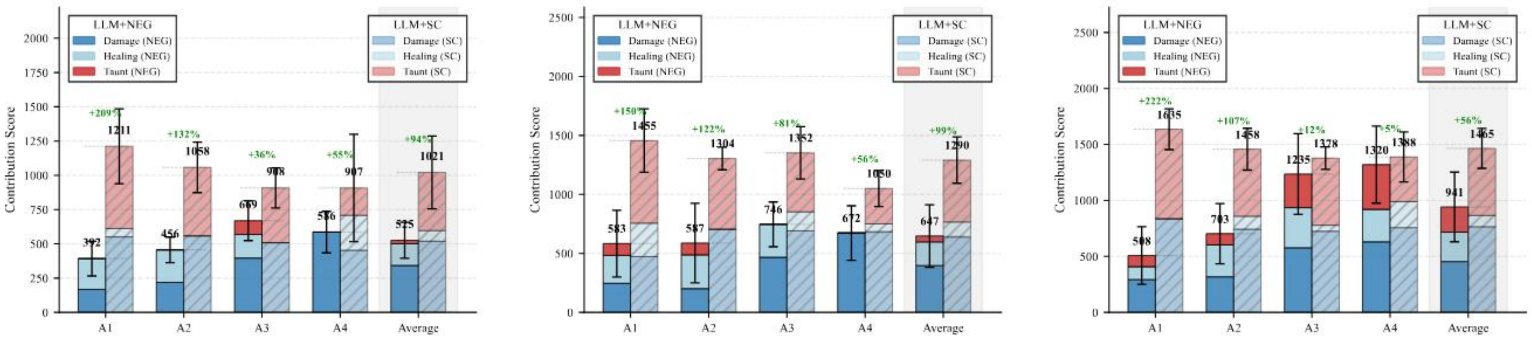


Figure 3: Comparison of agent payoffs and negotiation dynamics in the escape game. (a) illustrates the individual payoffs obtained under different methods. (b) presents the number of negotiation rounds and the resulting payoff differences using the ShapleyCoop workflow.

Experiment

Raid Battle: a multi-turn, multi-agent RPG scenario where four heroes must cooperate to defeat a boss while balancing self-interest and coordination challenges.



(a) Level 1 (b) Level 2 (c) Level 3

Figure 4: Comparison of Contributions for Raid Battle

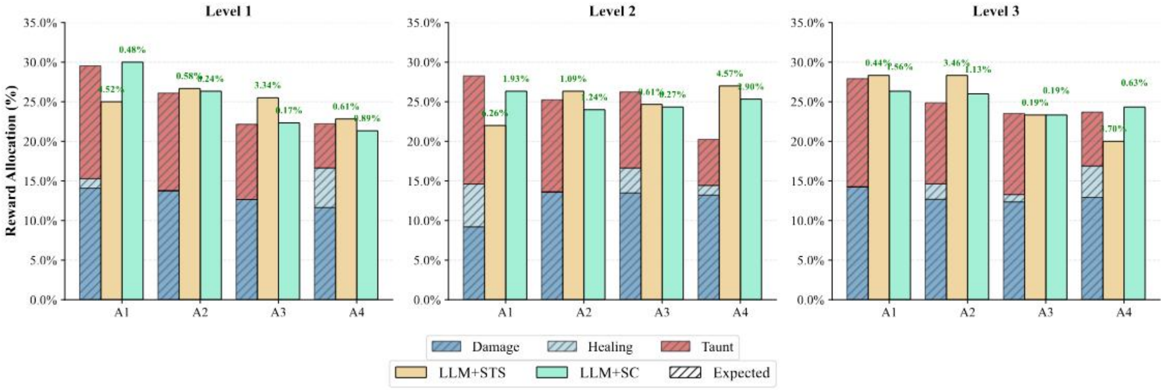


Figure 5: Comparison of Reward Allocation/Credit Assignment for Raid Battle

Experiment

ChatDEV: a virtual software company simulation where agents with defined roles (e.g., CEO, CTO, Programmer) collaborate through structured development tasks.

Table 3: Role contributions, allocated reward, and minimal adjustment

	BMI Calculator							ArtCanvas						
Role	**Code**	**Dec.**	**Docs**	**Fixes**	**WEV(%)**	**Reward(%)**	**Adj.(%)**	**Code**	**Dec.**	**Docs**	**Fixes**	**WEV(%)**	**Reward(%)**	**Adj.(%)**
CEO	0	3	0	0	7.5–17.5	15	0	0	2	0	0	4.3–10.0	5	0
Counselor	0	0	3	0	2.1–6.4	3	0	0	0	2	0	1.3–3.8	5	−1.3
CPO	0	1	4	0	5.4–14.4	20	−5.6	0	1	6	0	5.9–16.3	20	−3.8
CTO	0	2	0	0	5.0–11.7	25	−13.3	0	4	0	0	8.6–20.0	10	0
Programmer	45	0	0	3	30.9–47.1	25	+5.9	41	0	0	0	26.4–39.1	35	0
Reviewer	7	0	0	3	11.1–17.9	12	0	1	0	0	2	15.6–25.9	25	0

$$\mathrm{WEV}_r = \sum_{i \in \{\mathrm{code},\mathrm{dec},\mathrm{doc},\mathrm{fix}\}} \frac{\theta_{r,i}}{\sum_k \theta_{k,i}} w_i,$$

Standardized weights derived from benchmarks such as COCOMO II, COCOMO and CSBSG

(1) **BMI Calculator:** Develop an application calculating Body Mass Index from user inputs.
(2) **ArtCanvas:** Create a virtual painting studio app providing canvas, brushes, and color palettes.

Summary & Contributions

Pricing-Based Perspective for Multi-LLM Cooperation:

- Align heterogeneous goals of self-interested LLM agents using principled pricing mechanisms inspired by cooperative game theory, enabling spontaneous cooperation in open-ended scenarios.

Shapley-Coop: Cooperative Workflow:

- Introduce Shapley-Coop, combining Shapley Chain-of-Thought reasoning with structured negotiation protocols to ensure fair credit assignment, align incentives, and maintain agent autonomy.

Empirical Validation:

- Demonstrate robust cooperation, equitable reward allocation, and improved collaborative dynamics across social-dilemma and real-world software-engineering tasks, proving practical applicability and effectiveness.