

LitaAgentYR 谈判代理策略报告

LitaAgentYR Negotiation Agent Strategy Report

Lita, With assistance from Gemini

June 8, 2025

Contents

1	总体策略	3
1.1	策略支柱	3
2	实现与架构	4
2.1	关键组件与变量	4
2.2	核心方法调用流程	4
3	策略的数学建模	4
3.1	接受逻辑 (Acceptance Logic)	4
3.2	让步逻辑 (Concession Logic)	6
3.3	对手建模 (Opponent Modeling)	6
4	Overall Strategy	8
4.1	Strategic Pillars	8
5	Implementation and Architecture	9
5.1	Key Components and Variables	9
5.2	Core Method Call Flow	10
6	Mathematical Formulation of Strategy	10
6.1	Acceptance Logic	10
6.2	Concession Logic	11
6.3	Opponent Modeling	12

中文报告 (Chinese Report)

Abstract

本报告旨在深度剖析 **LitaAgentYR** 谈判代理的核心策略与算法实现。**LitaAgentYR** 是一个为供应链管理联盟 (SCML) 竞赛设计的、具有库存感知和动态风险管理能力的代理。报告首先阐述了其基于三层采购与产能约束销售的总体策略，随后详细说明了各个核心方法的功能、变量含义以及它们之间的协同工作流程。最后，报告使用数学公式对代理的接受准则、让步逻辑和对手建模算法进行了精确的形式化表达，并将这些模型与代码实现紧密结合，提供了从理论到实践的全面解析。

1 总体策略

LitaAgentYR 的核心策略是构建一个以**全局盈利和风险防范**为目标的长期主义代理。它不追求单次谈判的局部最优，而是通过紧密结合库存管理，动态适应市场变化，确保在整个模拟周期内的稳健表现。

1.1 策略支柱

1. **库存驱动决策**: 代理的核心大脑是其内部的 *InventoryManager* 模块。该模块采用“准时制 (Just-in-Time)”思想，通过反向传播销售订单来精确规划每日的生产任务，并由此计算出未来每日的原材料精确短缺量。所有谈判行为都以此数据为基础，实现了目标驱动的采购与销售。
2. **分层采购逻辑**: 为了平衡成本与风险，代理将采购需求细分为三个不同优先级的层次：
 - **紧急采购 (Emergency)**: 满足当日生产必需的原材料缺口，避免因生产停滞导致违约。在此场景下，代理的接受意愿最高，愿意以接近甚至略高于违约罚金的价格成交。
 - **计划采购 (Planned)**: 为满足未来已签订的销售订单而进行的远期采购。此决策的核心是在保证未来销售利润的前提下，权衡采购成本和仓储成本。
 - **机会性采购 (Optional)**: 当市场价格远低于历史均价时，在不显著增加库存积压风险的前提下进行战略性囤货，以降低整体原材料成本。
3. **产能约束销售**: 代理在对外销售时，严格遵守其每日的生产能力限制。它绝不签订任何在交货日期前无法完成生产的合同，从源头上杜绝了因无法履约而导致的巨额罚款。
4. **动态利润与风险管理**: 代理的利润目标和风险偏好是动态变化的。它会根据游戏阶段、库存健康度、以及历史销售转化率，实时调整其销售的最低利润率。

min_profit_ratio 和机会性采购的折扣阈值 $bargain_threshold$ 。

5. **帕累托改进的让步策略**: 在还价时, 代理不仅仅在价格维度上进行妥协。它会综合评估价格、数量和交货时间, 尝试提出对自己更有利, 同时又可能被对手接受的多维度组合方案, 沿着帕累托前沿探索双赢的可能。

2 实现与架构

2.1 关键组件与变量

LitaAgentYR 主类, 继承自 *StdSyncAgent*, 负责响应 SCML 环境的各种回调。

HeuristicSettings 一个 *dataclass*, 集中存储了代理所有的可调启发式参数, 如让步曲线指数、利润率、折扣阈值等, 增强了代码的可维护性和可配置性。

InventoryManager 库存管理模块, 是代理的决策中枢。负责追踪库存批次、管理待处理合同、生成生产计划以及预测物料短缺。

min_profit_ratio 销售时必须满足的最低利润率。

bargain_threshold 机会性采购的价格折扣阈值。

partner_models 字典, 为每个对手存储一个独立的在线逻辑回归模型参数 (w_0, w_1)。

2.2 核心方法调用流程

代理的谈判逻辑始于 *counter_all* 方法, 这是一个顶层分发器。

3 策略的数学建模

3.1 接受逻辑 (Acceptance Logic)

代理的接受决策是情境化的, 由四个不同模块的逻辑规则确定。

1. **紧急供应**: 接受来自供应商 i 的报价 (q_i, t_i, p_i) , 如果 t_i 是当天, 且满足:

$$\text{Accept if } (p_i \leq \pi) \vee (p_i \leq 1.1 \cdot \pi \wedge t_{\text{rel}} > 0.8)$$

其中 π 是违约罚金 (*shortfall_penalty*), t_{rel} 是相对谈判时间。

2. **计划性供应**: 接受报价需同时满足盈利性和库存容量限制。

- **盈利性**: $p_{\text{eff}} \leq p_{\text{max_affordable}}$

其中有效价格 $p_{\text{eff}} = p_i + C_{\text{storage}} \cdot (t_i - t_{\text{today}} - 1)$, 包含了仓储成本。可接受

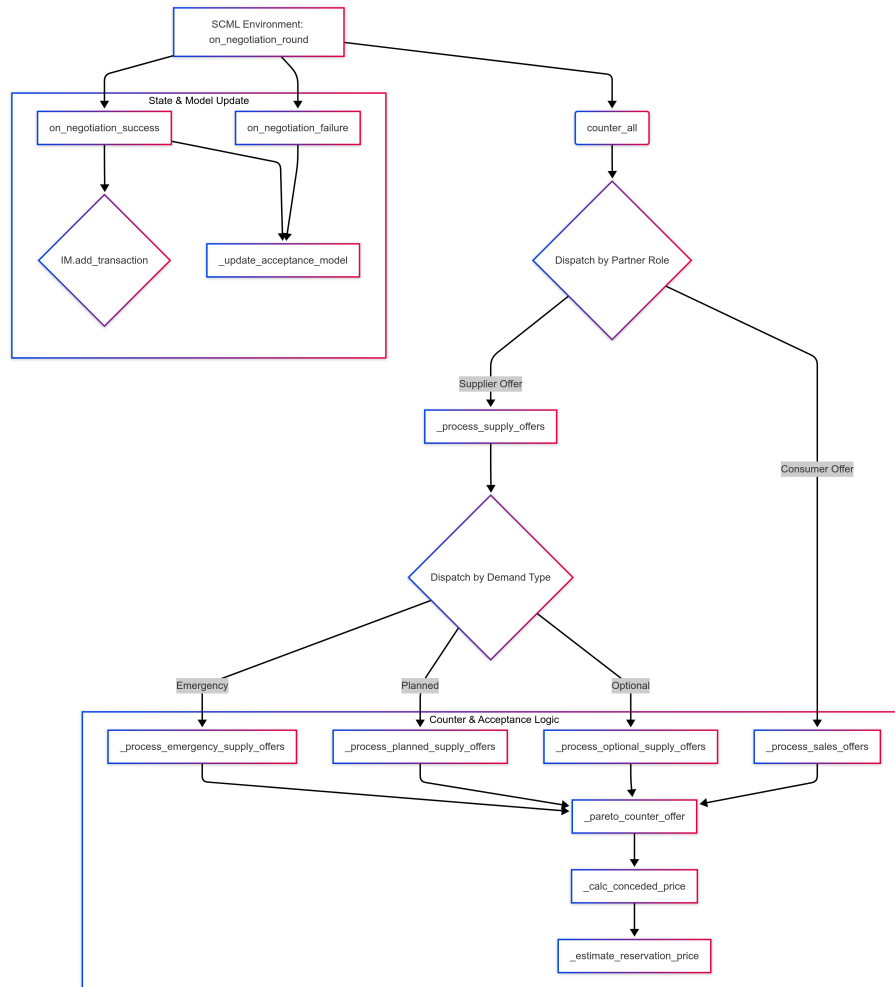


Figure 1: 核心方法调用流程

的最高采购价 $p_{\max_affordable}$ 由预估的未来产品售价 \hat{S} 、加工成本 C_{proc} 和最低利润率 m 决定：

$$p_{\max_affordable} = \hat{S} \cdot (1 - m) - C_{\text{proc}}$$

- **库存容量:** $q_i \leq H(t_i)$, $H(t_i)$ 是在 t_i 时刻的“采购空间”。

3. 机会性供应:

$$\text{Accept if } p_i \leq \bar{p}_{\text{market}} \cdot \delta_{\text{bargain}} \wedge q_i \leq H_{\text{opt}}(t_i)$$

其中 \bar{p}_{market} 是市场均价, δ_{bargain} 是折扣阈值 (‘bargain_threshold’)。

4. 销售:

$$\text{Accept if } q_i \leq C_{\text{avail}}(t_i) \wedge p_i \geq p_{\min_sell}$$

其中 $C_{\text{avail}}(t_i)$ 是在 t_i 时刻的可用产能, 最低售价 p_{\min_sell} 由平均原材料成本 $C_{\text{raw_avg}}$ 、加工成本 C_{proc} 和动态利润率 m_{dynamic} 决定：

$$p_{\min_sell} = (C_{\text{raw_avg}} + C_{\text{proc}}) \cdot (1 + m_{\text{dynamic}})$$

3.2 让步逻辑 (Concession Logic)

让步逻辑在 `_calc_conceded_price` 方法中实现, 分三步计算还价 p^{new} 。

1. 让步乘数 (β):

$$\beta(t_{\text{rel}}, r_{\text{opp}}) = \min(1, t_{\text{rel}}^{\gamma} \cdot (1 + r_{\text{opp}}))$$

其中, t_{rel} 是相对谈判时间, r_{opp} 是对手的相对让步率, γ 是让步曲线的幂 (concession_curve_power)。

2. 目标价格混合 (p^*):

$$p^* = \frac{p_{\text{agent_target}} + \hat{p}_{\text{opp}}}{2}$$

其中 $p_{\text{agent_target}}$ 是代理本轮的理想价格, \hat{p}_{opp} 是通过对手模型估算的对手保留价格。

3. 最终价格计算 (p^{new}):

$$p^{\text{new}} = p_0 + \beta \cdot (p^* - p_0)$$

其中 p_0 是代理的初始锚点价格 (对卖家是 NMI 最高价, 对买家是 NMI 最低价)。

3.3 对手建模 (Opponent Modeling)

代理为每个伙伴维护一个在线逻辑回归模型, 以预测其接受概率并推断保留价格。

1. **模型定义**: 接受概率 P_{acc} 被建模为价格 p 的 Sigmoid 函数:

$$P_{\text{acc}}(p) = \sigma(w_0 + w_1 \cdot x) = \frac{1}{1 + e^{-(w_0 + w_1 \cdot x)}}$$

特征 x 定义为: $x = p$ (采购时) 或 $x = -p$ (销售时)。

2. **模型更新** (`_update_acceptance_model`): 在每次交互后, 使用随机梯度下降更新权重 w_0, w_1 。设真实标签为 $y \in \{0, 1\}$, 预测概率为 $\hat{y} = P_{\text{acc}}(p)$, 则更新规则为:

$$w_0 \leftarrow w_0 + \eta \cdot (y - \hat{y})$$

$$w_1 \leftarrow w_1 + \eta \cdot (y - \hat{y}) \cdot x$$

其中 η 是学习率 (`h.logreg_learning_rate`)。

3. **保留价格估算** (`_estimate_reservation_price`): 假设对手的保留价格是其接受概率为 50% 的点, 此时 $w_0 + w_1 \cdot x_r = 0$ 。解出保留价格对应的特征值 $x_r = -w_0/w_1$, 并转换回价格 \hat{p}_{opp} 。这个估值是让步逻辑中混合目标价格的关键输入。

English Report

Abstract

This report provides an in-depth analysis of the core strategy and algorithmic implementation of the LitaAgentYR negotiation agent. Designed for the Supply Chain Management League (SCML) competition, LitaAgentYR is an agent with inventory awareness and dynamic risk management capabilities. The report first outlines its overall strategy, which is based on a three-tiered procurement system and capacity-constrained sales. It then details the functionality of its core methods, the meaning of key variables, and their collaborative workflow. Finally, the report presents a formal mathematical representation of the agent’s acceptance criteria, concession logic, and opponent modeling algorithms, closely linking these models to their code implementations to offer a comprehensive analysis from theory to practice.

4 Overall Strategy

The core strategy of LitaAgentYR is to build a long-term agent focused on global profitability and risk aversion. It does not pursue local optima in single negotiations but aims for robust performance throughout the entire simulation period by closely integrating with inventory management and dynamically adapting to market changes.

4.1 Strategic Pillars

1. **Inventory-Driven Decision Making:** The agent’s core brain is its internal InventoryManager module. This module employs a "Just-in-Time" philosophy, accurately planning daily production tasks by back-propagating sales orders and thereby calculating the precise daily shortfall of raw materials for the future. All negotiation activities are based on this data, enabling goal-driven procurement and sales.
2. **Tiered Procurement Logic:** To balance cost and risk, the agent categorizes procurement needs into three priority levels:
 - **Emergency Procurement:** Fulfills the immediate raw material gap required for the current day’s production, preventing production halts and contract breaches. In this scenario, the agent has the highest willingness to accept, even at prices close to or slightly above the shortfall penalty.
 - **Planned Procurement:** Procures materials for future production needs derived from signed sales orders. The core of this decision is to balance procure-

ment costs and storage costs while ensuring future sales profitability.

- **Optional Procurement:** When "bargains" appear on the market at prices significantly lower than the historical average, the agent engages in strategic stockpiling without significantly increasing inventory risks, aiming to lower the overall average cost of raw materials.
- 3. **Capacity-Constrained Sales:** When selling products, the agent strictly adheres to its daily production capacity. It never signs a contract that cannot be produced and delivered on time, fundamentally eliminating massive penalties from breach of contract.
- 4. **Dynamic Profit & Risk Management:** The agent's profit targets and risk appetite are not static. It dynamically adjusts its minimum sales profit margin `min_profit_ratio` and its willingness for opportunistic procurement `bargain_threshold` based on market conditions, game stage, inventory health, and historical sales conversion rates.
- 5. **Pareto-Aware Concession Strategy:** When making counter-offers, the agent does not just concede on price. It holistically evaluates price, quantity, and delivery time, attempting to propose multi-dimensional combinations that are more favorable to itself while potentially being more attractive to the opponent, thus exploring win-win possibilities along the Pareto frontier.

5 Implementation and Architecture

5.1 Key Components and Variables

LitaAgentYR The main class, inheriting from `StdSyncAgent`, responsible for responding to various callbacks from the SCML environment.

HeuristicSettings A dataclass that centralizes all tunable heuristic parameters of the agent, such as concession curve exponents, profit margins, and discount thresholds, enhancing code maintainability and configurability.

InventoryManager The inventory management module, serving as the agent's decision-making hub. It is responsible for tracking inventory batches, managing pending contracts, generating production plans, and forecasting material shortages.

min_profit_ratio The minimum profit margin that must be met when selling products.

bargain_threshold The price discount threshold for opportunistic procurement.

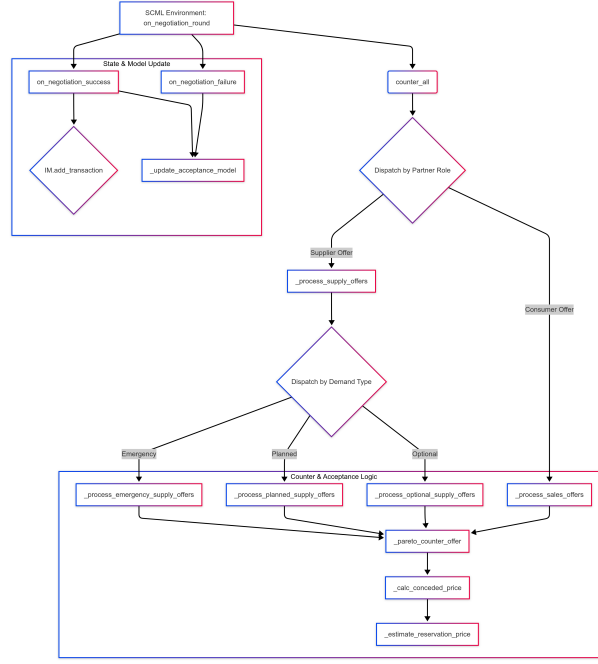


Figure 2: Core Method Call Flow

partner_models A dictionary storing a separate online logistic regression model (parameters w_0, w_1) for each opponent.

5.2 Core Method Call Flow

The agent’s negotiation logic begins with the **counter_all** method, which acts as a top-level dispatcher.

6 Mathematical Formulation of Strategy

6.1 Acceptance Logic

The agent’s acceptance decisions are contextual and determined by the logical rules within four different processing modules.

1. **Emergency Supply:** An offer (q_i, t_i, p_i) from supplier i is accepted if t_i is the current day and:

$$\text{Accept if } (p_i \leq \pi) \vee (p_i \leq 1.1 \cdot \pi \wedge t_{\text{rel}} > 0.8)$$

where π is the **shortfall_penalty** and t_{rel} is the relative negotiation time.

2. **Planned Supply:** Acceptance requires satisfying both profitability and inventory capacity constraints.

- **Profitability:** $p_{\text{eff}} \leq p_{\text{max_affordable}}$

where the effective price $p_{\text{eff}} = p_i + C_{\text{storage}} \cdot (t_i - t_{\text{today}} - 1)$ includes estimated storage costs. The maximum affordable price $p_{\text{max_affordable}}$ is determined by the estimated future product selling price \hat{S} , processing cost C_{proc} , and minimum profit margin m :

$$p_{\text{max_affordable}} = \hat{S} \cdot (1 - m) - C_{\text{proc}}$$

- **Inventory Capacity:** $q_i \leq H(t_i)$, where $H(t_i)$ is the procurement headroom at time t_i .

3. Optional Supply:

$$\text{Accept if } p_i \leq \bar{p}_{\text{market}} \cdot \delta_{\text{bargain}} \wedge q_i \leq H_{\text{opt}}(t_i)$$

where \bar{p}_{market} is the average market price and δ_{bargain} is the ‘bargain_threshold’.

4. Sales:

$$\text{Accept if } q_i \leq C_{\text{avail}}(t_i) \wedge p_i \geq p_{\text{min_sell}}$$

where $C_{\text{avail}}(t_i)$ is the available production capacity at time t_i , and the minimum selling price $p_{\text{min_sell}}$ is determined by costs and a dynamic profit margin m_{dynamic} :

$$p_{\text{min_sell}} = (C_{\text{raw_avg}} + C_{\text{proc}}) \cdot (1 + m_{\text{dynamic}})$$

6.2 Concession Logic

The concession logic, implemented in `_calc_conceded_price`, calculates a new counter-offer price p^{new} in three steps.

1. Concession Multiplier (β):

$$\beta(t_{\text{rel}}, r_{\text{opp}}) = \min(1, t_{\text{rel}}^\gamma \cdot (1 + r_{\text{opp}}))$$

where t_{rel} is relative time, r_{opp} is the opponent’s relative concession rate, and γ is the **concession_curve_power**.

2. Target Price Blending (p^*):

$$p^* = \frac{p_{\text{agent_target}} + \hat{p}_{\text{opp}}}{2}$$

where $p_{\text{agent_target}}$ is the agent’s ideal price for the round, and \hat{p}_{opp} is the opponent’s estimated reservation price from the opponent model.

3. Final Price Calculation (p^{new}):

$$p^{\text{new}} = p_0 + \beta \cdot (p^{\star} - p_0)$$

where p_0 is the agent's initial anchor price (NMI max for sellers, NMI min for buyers).

6.3 Opponent Modeling

The agent maintains an online logistic regression model for each partner to predict their acceptance probability and infer their reservation price.

1. **Model Definition:** The acceptance probability P_{acc} is modeled as a Sigmoid function of price p :

$$P_{\text{acc}}(p) = \sigma(w_0 + w_1 \cdot x) = \frac{1}{1 + e^{-(w_0 + w_1 \cdot x)}}$$

The feature x is defined as $x = p$ (for buying) or $x = -p$ (for selling).

2. **Model Update (`_update_acceptance_model`):** After each interaction, the weights w_0, w_1 are updated via stochastic gradient descent. With true label $y \in \{0, 1\}$ and prediction $\hat{y} = P_{\text{acc}}(p)$, the update rules are:

$$\begin{aligned} w_0 &\leftarrow w_0 + \eta \cdot (y - \hat{y}) \\ w_1 &\leftarrow w_1 + \eta \cdot (y - \hat{y}) \cdot x \end{aligned}$$

where η is the learning rate `h.logreg_learning_rate`.

3. **Reservation Price Estimation (`_estimate_reservation_price`):** The agent assumes the opponent's reservation price is where the acceptance probability is 50%, which implies $w_0 + w_1 \cdot x_r = 0$. This gives the reservation feature value $x_r = -w_0/w_1$, which is then converted back to a price \hat{p}_{opp} . This estimate is a key input for the concession logic's blended target price.