

# 扩展属性维度以容纳业务参数：状态泛化与策略编号机制

- 作者：GaoZheng
- 日期：2025-03-19
- 版本：v1.0.0

## 摘要

在实际交易建模中，策略参数（如下单间距、止损幅度等）对系统演化路径具有根本性影响。为了实现模型与策略的统一建模，需要将业务参数内生化为状态属性，形成“**策略参数嵌入状态编号**”的结构。这一机制使得每个状态不仅代表市场与账户的外部观测，还表示系统在具体策略参数配置下的演化倾向，从而实现状态空间、路径空间与策略空间的动态闭环耦合，并构成可解释的策略扰动反馈系统。特别地，引入“**转向机制**”允许系统在面临不良路径时探索参数扰动路径，例如：

$$\text{账户健康}_1 \Rightarrow \text{账户健康}_2$$

成为新路径的起点并带入样本空间进行训练积累。此方式最终构成**白盒强化学习结构的泛范畴表示**。

## 1. 状态编号机制的结构表达

引入策略参数集合：

$$\Theta = \{\theta_1, \theta_2, \dots\}$$

原始状态空间  $S = \{s_1, s_2, \dots\}$  扩展为：

$$\tilde{S} = S \times \Theta = \{(s_i, \theta_j)\}$$

每个编号状态  $(s_i, \theta_j)$  表示：状态  $s_i$  在参数配置  $\theta_j$  下的策略态。

例如：

- 账户健康<sub>1</sub> = (账户健康,  $\theta_1$ )
- 账户健康<sub>2</sub> = (账户健康,  $\theta_2$ )

这些泛化状态即是系统演化的策略—行为嵌套节点。

---

## 2. 路径空间的策略编号嵌套

系统不再考虑单一状态路径：

$$s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow \dots$$

而是复合态路径：

$$(s_1, \theta_1) \rightarrow (s_2, \theta_1) \rightarrow \dots \rightarrow (s_n, \theta_1)$$

当需要进行策略扰动时，路径将出现“编号转向”：

$$(s_k, \theta_1) \rightarrow (s_k, \theta_2) \quad \text{即：状态不变，策略参数转向}$$

该“转向态射”：

$$\tau : (s_k, \theta_1) \rightarrow (s_k, \theta_2)$$

成为策略扰动机制的基本操作路径，并带入新的窗口样本路径训练集中，例如：

$$(s_k, \theta_1) \rightarrow (s_k, \theta_2) \rightarrow (s_{k+1}, \theta_2) \rightarrow \dots$$

---

## 3. 不良路径应对与参数扰动机制

若观察到某一路径：

$$(s_1, \theta_1) \rightarrow (s_2, \theta_1) \rightarrow (s_3, \theta_1) \quad \text{为低得分}$$

系统可尝试路径扰动：

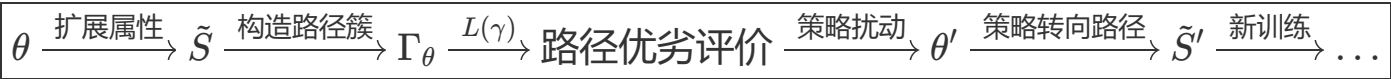
$$(s_1, \theta_1) \rightarrow (s_1, \theta_2) \rightarrow \dots \quad (\text{策略转向})$$

若该路径在现有样本库中缺失，则：

- 生成新策略路径；
- 使用路径积分模型模拟其得分；
- 若高于原路径则存储为“优化候选路径”；
- 否则记录无效扰动。

## 4. 路径演化与策略反馈结构图

构成如下反馈机制：



其中：

- $\tilde{S}$ ：策略-状态联合结构；
- $\Gamma_\theta$ ：策略下所有历史路径；
- $L(\gamma)$ ：路径积分；
- $\theta' \neq \theta$ ：新策略；
- $(s, \theta) \rightarrow (s, \theta')$ ：策略编号转向。

## 5. 学习与记忆：策略扰动的试探与积累

路径数据形式：

$$\{(\tilde{s}, \theta), \gamma_\theta, L(\gamma_\theta)\}$$

系统逐步构建一个基于  $(s, \theta)$  为索引的**策略-路径记忆结构**，用于：

- 最佳路径初始化；
- 策略扰动方向估计；
- 局部可解释性推理；
- 反向强化学习（路径  $\rightarrow$  参数）建模。

## 总结

该结构构成了如下闭环：



该结构具备：

- 策略扰动的路径可解释性；
- 状态粒度与策略空间的一体化建模；

- 支持结构性策略优化的闭环反馈系统；
- 为复杂决策系统提供白盒强化学习基础。

这是从交易行为建模迈向泛范畴可解释AI系统的重要路径。

---

## 许可声明 (License)

Copyright (C) 2025 GaoZheng

本文档采用[知识共享-署名-非商业性使用-禁止演绎 4.0 国际许可协议 \(CC BY-NC-ND 4.0\)](#)进行许可。