

# 仿真模拟 PRD

## 1. 模拟环境设置（沙盒设置）

### 1.1. 时间

环境中的时间分为阶段（stage）和帧（frame）。每个帧为一个完整行动（action）或互动（interaction）的最小时间单位，不同类型的行动或互动所需时间都为 1 帧。整个模拟过程的时间阶段用  $T_s$  表示，每个时间阶段包含的帧数用  $T_f$  表示。

批注 [Microsoft1]: 行动为个体内的行为，互动为个体间的行为。

### 1.2. 个体（Agent）

个体（Agent）是所有行为和互动的主体。模拟过程中设置 50 个 Agent。个体基于对环境信息的主观感知，进行行为决策，选定并执行行动或互动。个体间的互相影响由交互影响规则决定。

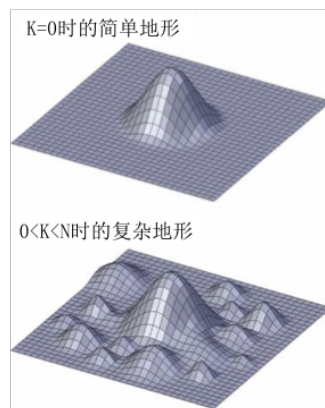
### 1.3. 行动方案与适应度地形

本研究采用 NK 模型来表示系统所在外环境的问题情景以及情景中所有可能的行动方案的结果。环境中个体所有可能的行动方案可以用一个包含 N 个组成成分的向量  $A(x_0, \dots, x_{N-1})$  来表示，其中每个组成成分  $x_i$  代表问题情景中影响行动方案的某因素状态。为简化模型令所有因素仅有 2 种状态，并用 0 和 1 来作为每个因素的状态标记，即  $x_i \in \{0,1\}$ 。此时，在一个问题情景中共存在  $2^N$  种可能的行动，构成一个行动空间。行动空间中任意一个行动方案可以用唯一的状态标记组合来表示，例如 N=4 时，存在行动方案  $A(1,0,1,1)$ 。任意一个行动方案的表现由结果函数  $F(x_0, \dots, x_{N-1})$  表示，具体公式如下：

$$F(x_0, \dots, x_{N-1}) = \frac{1}{N} \sum_{i=0}^{N-1} f_i(x_0, \dots, x_{N-1})$$

公式中  $f_i(x_0, \dots, x_{N-1})$  表示每个因素对行动方案结果的贡献，其结果依赖于每个因素自己的状态和另外 K 个相关因素的状态。在本模型中，每个因素与另外 K 个随机因素相关，组成一个邻接矩阵。在模型中，每一个因素对应的  $f_i(x_0, \dots, x_{N-1})$  都是满足  $U(0,1)$  均匀分布的随机数。结果函数  $F(x_0, \dots, x_{N-1})$  经过 min-max 归一化后映射到  $[0,1]$  区间，形成每个行动方案的适应分数  $\hat{F}(x_0, \dots, x_{N-1})$ 。如果将所有的潜在行动视作一个二维的行动空间，将每个行动方案的适应分数对应到行动空间中的相应位置，就构成了一张类似三维地形图适应度地形（fitness landscape）。地形中的每一个山峰就

代表一个有效的行动方案，山峰高度越高行动方案就越能解决外环境中的问题。根据 NK 模型原理，参数 N 的大小决定了行动空间大小，参数 K 决定了适应度地形的崎岖程度。当 K=0 时，所有行动方案的适应分数间彼此独立，适应度地形中仅有唯一的一处山峰，从任意起点出发最终都会收敛到唯一最有效方案上。而当 K 接近 N 时，地形变得非常崎岖，地图中山峰遍布，存在多种多样的有效行动方案，但每个方案的适应分数不尽相同。两种情况下的地形示意图如下所示：



#### 1.4. 适应度地形的动态变化与问题复杂度

在一次模拟过程中，适应度地形会动态变化。在变化过程中，每次适应度地形从初始状态到最终状态的变化时间为  $T_p$  个阶段。初始和最终状态的适应度地形都通过前述的 NK 模型生成。在一个变化时间内，每个行动方案在第  $T_i$  个阶段的适应分数为：

$$\hat{F}_i(x_0, \dots, x_{N-1}) = \hat{F}_{i-1}(x_0, \dots, x_{N-1}) + \frac{\hat{F}_p(x_0, \dots, x_{N-1}) - \hat{F}_0(x_0, \dots, x_{N-1})}{T_p}, \quad 1 \leq i \leq T_p$$

$\hat{F}_0(x_0, \dots, x_{N-1})$  表示初始阶段的适应分数， $\hat{F}_p(x_0, \dots, x_{N-1})$  表示最终阶段的适应分数。

若  $T_p$  大于等于  $T_s$ ，则地形变化随模拟程序时间的结束而停止。若  $T_p$  小于  $T_s$ ，则在整个模拟过程中存在多次适应度地形的变化，此时前一次变化中最终状态的适应度地形即作为后一次变化中初始状态的适应度地形。

问题情景中的问题复杂度由适应度地形参数 K，以及地形动态变化时间  $T_p$  所决定。当 K 较大时，适应度地形崎岖，较难确定有效的行动方案，而  $T_p$  较小时，适应度地形变化频率较高，行动方案结果更难预测。对于任意一次模拟过程，其问题情景的问题复杂度指标 CPLX 可以通过以下公式合成：

$$CPLX = 1 - 0.75^{\frac{T_s}{T_p}} \times \frac{1}{1 + e^{K-5}}$$

### 1.5. 行动目标与问题时间压力

在模拟环境中，行动空间中全部行动方案在 $p_a$ 百分位点的适应分数代表当前阶段环境对系统内个体的行动需求。当个体决定寻找行动方案来满足行动需求时，个体会基于行动需求形成主观的目标适应分数。个体的行动目标是在本阶段内寻找有效的行动方案，使其适应分数大于等于目标适应分数，具体行为策略和分数计算将在下文个体行为逻辑部分阐述。由于每个阶段包含 $T_f$ 帧时间，每个阶段个体寻找和调整行动方案的时间小于等于 $T_f$ 。因此 $T_f$ 数值较小时，反映出更高的问题时间压力。

## 2. 个体的行为逻辑

### 2.1. 个体差异参数

#### 2.1.1. 环境理解能力

个体的环境理解能力表示个体在多大程度上能够准确洞察环境中的信息，个体的环境理解能力参数 $k_{insight}$ 在区间(0,1)上服从正态分布， $k_{insight} \sim N(0.5, 0.5^2)$

#### 2.1.2. 创新意识

创新意识 $k_{inno}$ 表示个体多大程度上倾向于寻求新方法，取值范围[0.1, 0.9]

#### 2.1.3. 探索尝试意愿

探索尝试 $k_{try}$ 表示个体多大程度上倾向于换一种方法解决问题，取值范围[0.1, 0.9]

### 2.2. 个体的情景意识

个体对外环境问题情景的感知和理解可以表示为个体的情景意识。个体的情景意识包括已知的行动方案集合，感知到的问题复杂度和感知到的问题时间压力。

个体感知到的外环境问题复杂度为随机变量 $SA_{cplx}$ ，其概率密度函数在区间[0,1]上服从期望为CPLX，方差为 $e^{-\frac{4\pi K \times k_{insight}}{N}}$ 的正态分布，可记为：

$$SA_{cplx} \sim N\left(CPLX, e^{-\frac{4\pi K \times k_{insight}}{N}}\right)$$

## 问题复杂度

，情景意识主要由感知的环境挑战和备选行动集两部分组成。感知的环境挑战 PC 为对某一个当前环境挑战 C (t) 的主观感知，两者关系参考公式

$$PC(t) = C(t) + 1/k_{insight} * rand()$$

暂定为环境挑战加一个随机扰动，扰动大小受到个体环境理解能力倒数的加权影响。

备选行动集则是个体当前状态下所有可采取行动的坐标集合，以及对每个行动坐标在各个问题上的结果（如果知道结果的话）。

基本单元形式为[x, y, 状态, L<sub>1</sub>(x,y), L<sub>2</sub>(x,y), …… , L<sub>n</sub>(x,y)]。状态分为“未知”“已理解”“已探索”。当不知到 L (x, y) 时，全部用默认值 0.1 替代，状态为“未知”。

### 2.3. 个体对行动的评价

个体对任意一行动坐标 (x, y) 的评价 U (x, y) ，包括 (0, 0) Sensemaking, 符合以下公式：

V1.0 单人版先简化对行动的评价方式

若 (x, y) = (0, 0) ，即对于学习理解 Sensemaking 的效用评价固定为 0.5, U (x, y) =0.5,

其他情况下，

$$U(x, y) = \sum_{i=0}^n \frac{w_i \times L_i(x, y)}{n}$$

其中，w<sub>i</sub> 表示当前问题的重要程度，其公式为

$$w_i = \begin{cases} 0.5, & AS_i(t) - C_i(t) < -35 \\ 1, & -35 \leq AS_i(t) - C_i(t) < 3 \\ 0, & AS_i(t) - C_i(t) \geq 3 \end{cases}$$

### 2.4. 个体的行动动机

个体在每一帧时，是否采取行动

其行动动机 D (T<sub>i</sub>) 可以

## 2.5. 个体的行动模式

个体需要在学习理解 Sensemaking，进步提高 Exploitation，探索尝试 Exploration，3 类行动模式中进行抉择。

### 2.5.1. 学习理解 Sensemaking

学习理解指个体通过学习和理解环境，学习理解主要受到备选行动集中，“未知”行动坐标数量  $NUM_{unknown}$ 、已理解和已探索中  $U(x, y) > 0.5$  的坐标数量  $NUM_{good}$  影响。另外，个体的创新意识越高，基线越高。学习理解的概率权重  $ODDS_{sensemaking}$  满足公式

$$ODDS_{sensemaking} = 1 \times k_{inno} + \left( \frac{NUM_{unknown}}{k} + 0.5k - e^{0.5NUM_{good}} \right)$$

批注 [Microsof2]: 可能非常的不稳定，最好先局部调试下

学习理解在执行阶段的具体内容参考：

从探索起点行动集中随机获得 0-2k 个新行动坐标，加入备选行动集，用来表示思考新的行动解决问题。个体在本“帧”不执行任何行动坐标，不返回行动结果，也不改变适应状态 AS。

另外，可以有额外加入 Bonus，即随机获得 0-k 个备选集中有行动坐标但无结果的单元的结果值。用来表示对环境的越来越理解。

### 2.5.2. 进步提高 Exploitation

进步提高表示个体在同一类行动中不断练习，改善行动结果的表现。具有较高的基础概率，保证行动的连续性，同时受上一帧的行动评价影响。另外，个体的创新意识越低，基线越高。进步提高的概率权重  $ODDS_{exploitation}$  满足公式

$$ODDS_{exploitation} = \begin{cases} 0, & x = 0, y = 0 \\ 2 \times (1 - k_{inno}) + \tan \frac{\pi U(x, y)}{2}, & x > 0, y > 0 \end{cases}$$

批注 [Microsof3]: 回溯步长问题，是只看前一帧还是前 10 帧的平均。

进步提高在执行阶段的具体内容参考：

在上一帧行动坐标周围半径为  $l \times k_{insight}$  的区域内，随机获得一个相对最优的行动坐标，返回行动结果。将结果覆盖原备选行动集中上一帧行动的坐标所在单元。

### 2.5.3. 探索尝试 Exploration

探索尝试表示个体放弃当前行动，尝试新的行动,主要受到上一帧的行动评价影响。另外，个体的探索尝试意愿越高，越有可能切换。进步提高的概率权重  $ODDS_{exploration}$  满足公式

$$ODDS_{exploration} = 1 \times k_{try} - (1 - k_{try}) \tan \frac{\pi U(x,y)}{2}$$

探索尝试在执行阶段的具体内容参考：

在备选行动集中，个体选择除当前行动坐标和（0，0）外的一个新行动坐标，执行返回行动结果。所有备选行动坐标的权重为行动评价  $U(x,y)$ ，随机选择。

批注 [Microsof4]：回溯步长问题，是只看前一帧还是前 10 帧的平均。

2.6. 个体的行动逻辑

个体需要选择三种行动模式中的一种，只取 3 种行动模式中  $ODDS>0$  的行动模式进行选择。每种行动模式被选中的概率为  $ODDS/SUM\_ODDS$ 。（你说的纸带比喻）

2.7. 行动结果的执行与分配

个体在  $t$  时间帧上执行的行动坐标，对每一个问题都会返回结果。结果更新至  $AS(t)$  上。

2.8. 反馈

个体在过程中不改变个体差异函数。

3. 输出原始数据表

3.1. 环境挑战函数表

各个问题的环境挑战在每一帧的环境挑战数值和感受到的挑战数值

AgentID	时间 (帧)	$C_1(t)$	.....	$C_n(t)$	$PC_1(t)$	.....	$PC_n(t)$
1	1						
1	2						
1	.....						
1	10000						

3.2. 个体参数

虽然单人版只有一条记录，也先输出吧

AgentID	环境理解能力	创新意识	探索尝试意愿	.....
1				

