低资源消耗的AI人格模型，用于沙盒游戏或者任何可以实现 真人-AI 或者 AI-AI 交流的交流模拟。

每个智能作为一个包装后的自动机存在，拥有一套通用输入(感官)和输出（行为）。通过输入来获得对自身状况，环境状况，输入信息（交流文本或图片）的 记忆和认识。通过输出来控制搭载的人物模型来改变自身状况，改变环境，输出信息。输入和输出方式根据人物定义来裁剪，比如聋子，哑巴，瞎子，失去双腿……。

上述内容事实上是人格模型搭载到人物模型后与外界交互的问题。下面，我将描述人格模型内部的思考和决策。1. 任何人格的基础都是知识或者称记忆。2. 知识的不同是人格不同的最主要原因。3. 知识的最小单元为实体（物理的或概念的，比如石头和颜色）和事件（什么时候在哪里发生了什么，什么时候某人在哪里做什么）和关系（很多彼此有某种关系的石头组成房子， 某人掌握某种关系的知识可以建造房子）。4. 知识会随事件或时间丢失（越早的记忆缺损越多，越重要的记忆记得越牢）。5. 知识可以输出到某个载体，语言。绘画或抽象了绘画产生的文本。6. 可以通过经历事件或者交流获得或创造新的知识。7. 知识可以被事件或时间或别的新知识或者联想模糊化和扭曲（知识中的元素损失为更概略的上级元素，或者被错误的细化（臆测）为另一种相似的元素），包括任何实体，概念，事件和关系。8. 知识的储存包括长期和短期两种， 在调用短期知识失败时会试图调用长期知识。9. 每一次对知识的调用都会触发知识损失事件。但是经常被调用且损失率低的知识会更有机率逃脱或减少知识损失。10. 记忆属性好的人物模型，短期知识库更大且知识损失的机率更低。聪明的人物模型，可以在一次知识调用中调动更多层的相关知识并发现不自洽的知识。好奇的人物模型会经常反刍知识库找到知识之间可能的的联系以此来修正不自洽的知识。谨慎的人物模型选择多次验证后再修改可行率低的知识并在活动中选择可信度更高的知识。冲动的人物模型选择在决策中更加忽视或将可信度的重要性降低。

这里有个真正的问题，人格模型必须在一定程度上会执行自我欺骗和主动忽略一些可信度高的知识。

一些反复被验证为高可信度的知识可以保存在特殊记忆区，组成人格的“行为模式”（可抽象为三观），单位时间内行为模式成功应对率越高的总可信度越高。其中通过反刍知识抽象出一些当前记忆区最基础的关系和模式，可以命名为“信仰”。信仰和三观总可信度越高的，心理自信越强（排除物理环境和人物物理能力等外界因素 ）。行为模式可以有几套根据自我判断选择使用。当然这些内容也和一般知识一样需要反思修改保持自洽，但是对这部分的反思会更频繁。当有外界输入或者反思结果会涉及“信仰和三观”层的正确性和修改时，会受到特殊记忆区总可信度的反对。总可信度越大，反对性越强。修改幅度越大，反对性越强。会导致几种结果：修改“信仰和三观”层的知识，造成长期的知识反刍和重新建立自洽和可信度。略微修改，后果小得多。不修改，保留修改建议在短期记忆区，并标记不可信，等待反刍和自洽证明。不修改，将修改建议归入长期记忆，并对推导出修改建议的相关知识标记不可信，等待反刍和自洽证明。对修改建议的反刍和自洽证明优先级取决于此建议的重要性和可靠性。另附说明：短期记忆区的记忆不会真的丢失，而是转入长期记忆区接受模糊化和丢失的处理。思考过程中的相关记忆会调入短期记忆区，随时间自然沉降最后进入长期记忆区。

至于人格的推导和匹配的模式，我们引入经典心理学的三重人格理论。本我，自我，和超我。分别以现实利弊，社会道德，本能存续这三种基础行为原则对事件分析和判断，最后由“自我”来根据三种人格各自的行为计划推演结果进行反刍并选择最终本轮反应的行动方案。方案的执行有长期性，这部分知识放置于“计划”特殊记忆区，与长期记忆区有相同的遗忘和丢失逻辑。

一些补充：“本我，自我，和超我。分别以现实利弊，社会道德，本能存续这三种基础行为原则对事件分析和判断” 事实上，在本人格模型里，本我是以本个体的物理存续为原则的，是真实的人类人格中最基础的部分。超我在人类人格中是最上层建筑，在本模型中主要依托于“信仰和三观”区，以对“信仰和三观”的稳定性的保护，和“与其它对自己曾经总体有益的对象（任何对象，物体，个人， 组织， 关系）”的保护，为原则；亦即行为模式的不变性和对所在的某个关系群的生存有益。自我，则扮演着独立角色和三重人格决策博弈的仲裁角色。 三重人格决策博弈的仲裁过程至少有一下几轮： 0. 将激励起本轮博弈的输入信息和事件存入短期记忆区。 1. 自我，本我，超我，各自调用短期记忆挑选行为模式做出第一轮决策，存入各自的临时计划区。 2. 根据三种决策中涉及的关系，对象，资源消耗，目标效果，成功概率，分别推演预测结果分支得分和概率。 3. 将各种结果分支博弈，得分高且成功率高者胜。 第一轮获胜决策及推演分支分别存入短期记忆区。 4. 如果上一轮获胜决策评分很低，或第一第二得分的决策没有明显差别，进入反思阶段。否则，将最终决策放入计划区。 5. 自我，本我，超我，各自调用所有记忆区的较新的记忆重新选择的行为模式做出下一轮决策。 6. 重复2~4，但是记忆区可以调用更古老的记忆。 7. 根据人物的聪明程度，越聪明则允许每次调用记忆区长度和反思次数上限越长。 8. 达到反思次数上限或在4中分数过低时，执行默认计划：不变更当前行为，调节感官得到更多信息输入，执行计划列表中已在执行的计划。 所有的限制性参数，除了默认值外均不得写为定值，必须是随历史计算和变更后的当前值。

## 低资源消耗的AI人格模型设计与实现（由用户和AI共同提供）

**摘要（由AI提供）**

本文提出了一种适用于沙盒游戏及多智能体交流模拟的低资源消耗AI人格模型。该模型以自动机为基础，结合知识记忆、行为决策、三重人格理论等心理学原理，实现了高效、可扩展的人格建模。论文详细阐述模型的知识结构、记忆机制、决策流程，并给出实验环境与基础架构建议，旨在为AI角色扮演、虚拟人交互等领域提供理论与实践参考。

**1. 引言（由AI提供）**

随着AI在游戏、虚拟世界和人机交互中的广泛应用，如何在有限资源下实现真实、复杂的人格行为成为重要课题。传统AI人格模型多依赖大规模计算与数据，难以在资源受限环境下部署。本文基于用户提出的知识结构和决策逻辑，设计了一套低资源消耗的AI人格模型，兼顾效率与表现力。

**2. 人格模型总体架构（由用户提供）**

每个智能体以自动机形式存在，拥有通用输入（感官）和输出（行为）接口。输入用于感知自身状态、环境状况及交流信息（文本或图片），输出用于控制角色模型，实现状态变更、环境影响及信息输出。输入输出方式可根据角色定义裁剪，如聋哑、失明、肢体残疾等特殊情况。低资源消耗的AI人格模型.docx

**3. 知识与记忆系统（由用户提供）**

**3.1 知识结构（由用户提供）**

知识分为实体（物理或概念，如石头、颜色）、事件（时间、地点、行为）、关系（实体间的联系，如房屋结构、技能掌握）。知识是人格的基础，不同知识结构决定不同人格。低资源消耗的AI人格模型.docx

**3.2 记忆机制（由用户提供）**

知识随事件或时间丢失，早期记忆损失较多，重要记忆保存更牢。知识可通过语言、绘画等方式输出，也可通过事件或交流获得或创造。知识会因事件、时间或新知识联想而模糊或扭曲，包括实体、概念、事件和关系。低资源消耗的AI人格模型.docx

**3.3 短期与长期记忆（由用户提供）**

知识分为短期和长期记忆。调用短期知识失败时尝试调用长期知识。每次调用知识会触发知识损失事件，但频繁调用且损失率低的知识更易保留。记忆属性好的角色短期知识库更大，损失率更低；聪明角色能调动更多层相关知识并发现不自洽知识；好奇角色常反刍知识库修正知识；谨慎角色多次验证后才修改知识；冲动角色决策时降低可信度权重。低资源消耗的AI人格模型.docx

**4. 行为模式与信仰系统（由用户提供）**

反复验证为高可信度的知识可保存在特殊记忆区，组成人格的“行为模式”（三观）。行为模式成功应对率越高，总可信度越高。通过反刍知识抽象出基础关系和模式，形成“信仰”。信仰和三观总可信度越高，心理自信越强。行为模式可多套切换，并需频繁反思修正以保持自洽。修改信仰和三观时，特殊记忆区总可信度越大，反对性越强，修改幅度越大，反对性越强。修改建议可进入短期或长期记忆区，等待反刍和自洽证明，优先级取决于建议的重要性和可靠性。低资源消耗的AI人格模型.docx

**5. 记忆区转移与遗忘机制（由用户提供）**

短期记忆不会真正丢失，而是转入长期记忆区，接受模糊化和丢失处理。思考过程中的相关记忆会调入短期记忆区，随时间沉降最终进入长期记忆区。低资源消耗的AI人格模型.docx

**6. 三重人格决策博弈（由用户提供）**

**6.1 理论基础（由用户提供）**

引入经典心理学三重人格理论：本我、自我、超我。分别以现实利弊、社会道德、本能存续为基础行为原则，对事件进行分析和判断，最终由自我仲裁选择本轮反应方案。低资源消耗的AI人格模型.docx

**6.2 决策流程（由用户提供）**

1. 输入信息和事件存入短期记忆区。
2. 本我、自我、超我各自调用短期记忆，挑选行为模式做出第一轮决策，存入临时计划区。
3. 根据决策涉及的关系、对象、资源消耗、目标效果、成功概率，推演结果分支得分和概率。
4. 各分支博弈，得分高且成功率高者胜，获胜决策及分支存入短期记忆区。
5. 若获胜决策评分低或无明显差别，进入反思阶段，否则将最终决策放入计划区。
6. 三重人格各自调用所有记忆区较新记忆，重新选择行为模式做下一轮决策。
7. 重复推演与博弈，允许调用更古老记忆，聪明角色可增加调用长度和反思次数。
8. 达到反思次数上限或评分过低时，执行默认计划：不变更当前行为，调节感官获取更多信息，执行已在计划列表中的任务。
9. 所有限制性参数均随历史计算和变更动态调整。低资源消耗的AI人格模型.docx

**7. 模型实现建议（由AI提供）**

**7.1 数据结构设计**

* 实体、事件、关系采用图结构存储，节点为实体/事件，边为关系。
* 记忆区采用分层缓存机制，短期区高频访问，长期区低频访问并定期模糊化。
* 行为模式与信仰区采用权重标记，动态调整可信度。

**7.2 决策算法**

* 三重人格决策采用多轮博弈推演，评分函数结合资源消耗、目标达成率、社会道德等多维度。
* 反思机制通过知识库遍历与自洽性检验，自动修正不一致知识。

**7.3 资源优化**

* 输入输出裁剪，按角色能力动态调整感官与行为接口。
* 记忆损失与模糊化采用概率模型，降低存储与计算压力。
* 决策流程采用异步与并行处理，提升响应效率。

**8. 实验环境与基础架构建议（由AI提供）**

* 推荐使用Python或C++实现原型，结合网络图数据库（如Neo4j）管理知识结构。
* 记忆区可用Redis等内存数据库实现分层缓存。
* 行为决策模块建议采用多线程或协程优化。
* 沙盒环境可选用Unity或Unreal Engine集成AI角色，便于测试交互与行为表现。
* 实验建议包括：不同记忆损失率、决策轮数、角色属性对行为表现的影响对比；多智能体交流场景下的资源消耗与响应速度评测。

**9. 结论（由AI提供）**

本文系统提出并扩展了低资源消耗AI人格模型，结合知识结构、记忆机制、三重人格决策等多维度设计，兼顾理论深度与工程实现。模型适用于沙盒游戏、虚拟人交互等多场景，具备良好扩展性与资源效率。

**参考文献（由AI提供）**

1. Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach. Pearson.
2. Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction. MIT Press.
3. Freud, S. (1923). The Ego and the Id. W. W. Norton & Company.
4. McClelland, J. L., & Rumelhart, D. E. (1986). Parallel Distributed Processing. MIT Press.
5. Neo4j Graph Database Documentation. https://neo4j.com/docs/
6. Redis Documentation. https://redis.io/documentation/

根据AI模型的泛用性需求，Agent在环境中训练但必须与环境（及其他Agent）解耦，输入输出均应独立且不绑定特定训练环境。细化工程实施内容，详细描述训练初始条件（AI和环境）、训练步骤和方法，并生成完整的《低资源消耗AI人格模型的工程实施方案》。

## 低资源消耗AI人格模型的工程实施方案（由AI提供）

**1. 总体架构与参数设计（由AI提供）**

**1.1 主要参数**

* **记忆容量参数**：短期记忆容量、长期记忆容量、知识损失率、记忆反刍频率。
* **人格属性参数**：好奇心、谨慎度、冲动性、聪明度、行为模式可信度、信仰强度。
* **输入输出参数**：感官类型（视觉、听觉、触觉等）、行为接口（动作、语言、绘画等）。
* **决策参数**：三重人格权重、本我/自我/超我决策优先级、反思轮数上限、计划区容量。
* **环境接口参数**：环境状态描述、事件队列、Agent输入输出缓冲区。

**1.2 配置建议**

所有参数应支持动态调整，允许根据历史行为和环境反馈自适应优化。Agent的输入输出接口应为抽象化的数据结构，与具体环境实现完全解耦。

**2. 类实体与方法定义（由AI提供）**

**2.1 类实体设计**

- Agent（智能体）

- 属性：Memory（记忆系统）、Personality（人格体系）、Sensor（感官）、Actuator（行为器）、BeliefSystem（信仰系统）、Plan（计划区）、InputBuffer（输入缓冲区）、OutputBuffer（输出缓冲区）

- 方法：perceive()、think()、act()、reflect()、learn()、updateMemory()

- Memory（记忆系统）

- 属性：ShortTermMemory、LongTermMemory、KnowledgeLossRate

- 方法：store()、retrieve()、forget()、consolidate()

- Personality（人格体系）

- 属性：Curiosity、Cautiousness、Impulsiveness、Intelligence、BehaviorPatterns、Beliefs

- 方法：selectBehavior()、evaluateConsistency()、updateBelief()

- Sensor（感官）

- 属性：类型（视觉、听觉等）、输入接口

- 方法：sense(input\_data)

- Actuator（行为器）

- 属性：类型（动作、语言等）、输出接口

- 方法：execute(output\_data)

- BeliefSystem（信仰系统）

- 属性：Beliefs、Credibility

- 方法：updateBelief()、evaluateResistance()

- Plan（计划区）

- 属性：CurrentPlans、PlanHistory

- 方法：addPlan()、removePlan()、evaluatePlan()

- EnvironmentInterface（环境接口）

- 属性：StateDescription、EventQueue

- 方法：getState()、sendEvent()、receiveEvent()

**2.2 方法说明**

* **perceive()**：从InputBuffer采集信息，存入短期记忆。
* **think()**：启动决策线程，调用三重人格决策博弈，生成行为方案。
* **act()**：根据决策结果，将行为输出到OutputBuffer。
* **reflect()**：对行为结果进行反思，修正知识和信仰。
* **learn()**：通过事件或交流获得新知识，更新记忆系统。
* **updateMemory()**：定期进行记忆损失和模糊化处理。
* **sense(input\_data)**：处理抽象输入数据，不依赖具体环境。
* **execute(output\_data)**：生成抽象输出数据，供环境或其他Agent消费。

**3. 类之间的调用关系（由AI提供）**

* Agent通过Sensor采集InputBuffer中的信息，调用Memory存储感知结果。
* Agent在think()时，调用Personality进行三重人格决策，涉及BeliefSystem评估信仰阻力。
* Personality根据Memory中的知识和行为模式，选择行为方案，Plan区存储当前计划。
* Agent通过Actuator将行为输出到OutputBuffer，行为结果反馈至Memory和Personality。
* Agent周期性调用reflect()，对行为和信仰进行反思和修正。
* learn()方法在Agent与外界交流或经历事件时触发，更新Memory和BeliefSystem。
* EnvironmentInterface仅负责数据交换，不参与Agent内部逻辑。

**4. 多模输入输出的实现（由AI提供）**

**4.1 输入实现**

* Sensor基类支持多种感官（视觉、听觉、触觉等），每种感官实现sense(input\_data)方法，采集抽象化数据。
* 输入数据统一转化为知识单元，存入Memory。
* InputBuffer为标准化数据结构（如JSON、张量等），与环境解耦。

**4.2 输出实现**

* Actuator基类支持多种行为输出（动作、语言、绘画等），每种行为实现execute(output\_data)方法。
* OutputBuffer为标准化数据结构，供环境或其他Agent消费。
* 输出接口根据角色能力裁剪，如聋哑、失明等特殊情况屏蔽对应输入输出。

**5. 设计模式与应用点（由AI提供）**

**5.1 观察者模式**

* Sensor与Memory之间采用观察者模式，感官变化自动通知记忆系统更新。

**5.2 策略模式**

* Personality中的行为模式选择采用策略模式，支持多套行为方案动态切换。

**5.3 状态模式**

* Agent的行为和决策流程采用状态模式，区分感知、思考、行动、反思等不同阶段。

**5.4 责任链模式**

* 三重人格决策博弈采用责任链模式，依次由本我、自我、超我进行决策和仲裁。

**5.5 适配器模式**

* Sensor和Actuator通过适配器模式与环境接口对接，保证输入输出的抽象化和解耦。

**6. 思考线程的启动条件与参数（由AI提供）**

**6.1 启动条件**

* InputBuffer收到新数据或事件。
* 计划区任务到达执行时间或需重新评估。
* Memory中知识自洽性降低，触发反思。

**6.2 线程参数**

* 反思轮数上限、记忆调用深度、人格属性权重、决策超时时间。

**7. 线程调度与协作（由AI提供）**

* 每个Agent可独立启动思考线程，主线程负责感知与行为，子线程负责决策与反思。
* 多Agent场景下，采用消息队列或事件总线实现线程间协作与同步。
* 线程调度优先级根据事件紧急性、计划区任务优先级动态调整。
* 反思线程与决策线程可并行运行，反思结果实时反馈至决策流程。
* 环境仅作为数据交换平台，不干涉Agent内部逻辑。

**8. 新生状态知识与人格体系的准备与训练（由AI提供）**

**8.1 基础知识准备**

* 预设实体、事件、关系三类基础知识，如常见物体、基本动作、常用交流语句。
* 设定初始信仰和三观，赋予角色基础行为模式和信仰区可信度。
* 所有知识以抽象数据结构存储，避免与具体环境绑定。

**8.2 人格体系训练方案**

* 通过模拟事件和交流，逐步丰富知识库和行为模式。
* 采用有监督学习或强化学习方法，调整人格属性参数，使角色行为更贴合预期。
* 训练过程中，定期进行知识反刍和自洽性检验，修正不一致知识和信仰。
* 多角色协作训练，模拟AI-AI和AI-真人交流场景，提升模型泛化能力。
* 训练环境通过EnvironmentInterface提供抽象化输入输出，保证Agent与环境解耦。

**9. 训练初始条件与步骤（由AI提供）**

**9.1 初始条件**

* **AI Agent初始化**：加载基础知识库、人格参数、行为模式、信仰体系。
* **环境初始化**：通过EnvironmentInterface定义抽象化状态描述和事件队列，不包含具体物理细节。
* **输入输出缓冲区初始化**：Agent与环境通过标准化数据结构进行交互。

**9.2 训练步骤**

1. **环境生成抽象事件或状态描述**，写入Agent的InputBuffer。
2. **Agent感知输入**，将数据转化为知识单元，存入Memory。
3. **Agent启动思考线程**，进行三重人格决策博弈，生成行为方案。
4. **Agent将行为输出到OutputBuffer**，环境或其他Agent读取并反馈结果。
5. **Agent周期性反思与学习**，修正知识库和信仰体系。
6. **多Agent协作训练**，通过事件总线或消息队列实现信息交换，保证各Agent独立运行。
7. **定期评估Agent行为表现**，调整参数，优化泛化能力。

**9.3 训练方法**

* **无监督训练**：环境随机生成事件，Agent自主学习和调整行为。
* **有监督训练**：人为设定目标行为，环境反馈奖励或惩罚，Agent调整决策逻辑。
* **强化学习**：Agent根据环境反馈优化行为策略，提升任务完成率和自洽性。
* **迁移学习**：在不同环境下训练Agent，验证其泛化能力和解耦效果。

**10. 实施建议与总结（由AI提供）**

* 推荐采用面向对象编程语言（如Python、C++、Java）实现各类实体与方法。
* 多模输入输出建议结合深度学习模型（如CNN、RNN）处理复杂感官数据。
* 线程调度可用协程或多线程库（如Python asyncio、C++ std::thread）。
* 多Agent协作建议引入分布式消息队列（如RabbitMQ、Kafka）。
* 训练阶段可结合仿真平台（如Unity、Unreal Engine）进行场景模拟与评测。
* 所有输入输出均采用抽象化数据结构，保证Agent与环境和其他Agent完全解耦，提升模型泛用性和可移植性。

## 低资源消耗AI人格模型引入机器学习低资源经典模型的场景

在不使用大模型的前提下，如何在低资源消耗AI人格模型的工程实现中合理引入经典RNN、LSTM等较简单的神经网络模型进行简化和优化，结合模型架构和工程实施方案，从**知识处理、记忆机制、行为决策、多模输入输出**等关键环节，分析RNN/LSTM的适用场景，和具体建议。

**1. 知识与记忆系统中的应用**

**1.1 序列化知识处理**

* **场景**：知识的获取、记忆和遗忘过程本质上是时序数据的处理。例如，Agent在与环境或其他Agent交互时，输入的信息往往是连续的文本、事件流或感官数据流。
* **建议**：可用RNN/LSTM对输入的事件序列进行编码，捕捉时间上的依赖关系，实现知识的动态更新和遗忘。例如，短期记忆区可用LSTM单元存储最近的交流内容，长期记忆区则用更稀疏的RNN结构定期归档和模糊化知识。

**1.2 记忆损失与模糊化**

* **场景**：知识随时间和事件发生损失或模糊化，需模拟记忆的衰减和泛化过程。
* **建议**：利用RNN/LSTM的遗忘门机制，自动调整知识单元的保留与丢失概率，实现记忆的自然沉降和模糊化。

**2. 行为决策与反思机制中的应用**

**2.1 行为序列生成**

* **场景**：Agent在决策时，需根据历史行为和当前状态生成合理的行为序列。
* **建议**：采用LSTM生成行为序列，结合三重人格决策博弈，将本我、自我、超我的决策过程转化为时序决策问题，提升行为模式的连贯性和合理性。

**2.2 反思与自洽性检验**

* **场景**：Agent需周期性反思知识库，检验并修正不一致的知识和信仰。
* **建议**：用RNN/LSTM遍历知识序列，自动检测和标记自洽性低的知识片段，辅助Agent进行知识反刍和信仰修正。

**3. 多模输入输出的处理**

**3.1 序列化感官数据**

* **场景**：视觉、听觉等感官输入往往是时序数据（如语音、视频帧）。
* **建议**：用RNN/LSTM对多模态输入进行序列建模，实现对复杂感官信息的抽象化处理。例如，语音输入可用LSTM进行特征提取，视觉输入可结合CNN+LSTM进行时序分析。

**3.2 行为输出序列化**

* **场景**：Agent输出的动作、语言等行为也具有时序特性。
* **建议**：用LSTM生成动作序列或对话文本，提升输出的连贯性和上下文相关性。

**4. 训练与泛化能力提升**

**4.1 迁移学习与环境解耦**

* **场景**：Agent需在不同环境下训练并保持输入输出的抽象化。
* **建议**：用RNN/LSTM作为知识和行为的底层编码器，将环境特定信息转化为统一的抽象表示，便于迁移和泛化。

**4.2 强化学习结合**

* **场景**：Agent通过与环境交互不断优化行为策略。
* **建议**：将RNN/LSTM与强化学习算法结合，作为策略网络或价值网络的基础，提升决策的时序关联性和长期规划能力。

**总结与建议**

* RNN/LSTM适合处理模型中的**时序数据**，如知识流、事件流、行为序列、感官输入等。
* 可用于**记忆机制、行为决策、反思修正、多模输入输出**等环节，提升模型的动态性和泛化能力。
* 结合强化学习和迁移学习，可进一步增强Agent的环境适应性和解耦能力。
* 由于RNN/LSTM资源消耗远低于大模型，适合在低资源场景下实现高效的AI人格建模。