**会议纪要**

**会议主题：基于大语言模型的智能体技术**

**会议时间：** 2025年7月14日19：00至21：00

**会议地点：** 西安电子科技大学北校区会议中心203会议室

**线下人员：** 承楠 孙瑞锦 尹志胜 周淙浩 贺靖超 王兆薇 王葳 王秀程 周新阳 胡陆莹 齐阁 邓川 方忠盛 贾宏刚 郑雯馨 祝馨平 贾昊燏 张智杰 邱子仪 刘永红 沈京龙 韩松明 郑佩林 孙路路 王天宇 于凡迪 尉家豪 杨双宇 傅连浩 侯毓真 李青壮 闵昕阳 冉艺泉 刘琪 张岳 雷雨欣 曹江凌 龚臣 吴娅兰 李成成 陈梦豪

**线上人员：** 尚佳瑶 赵璇 郭译凡 许嘉洁 魏子超 朱军 解思舀 李昊坤 康贺延

**记 录 人：**杨双宇

**一.会议内容**

* **郭译凡：**传统的Agent存在一个问题，大部分的环境设置都是在一个虚拟的环境中训练的，很难在实际应用中估计这个Agent的误差。然后，如果大部分的实验都是基于实际行动的话，成本又太高了。因此我们引出了今天介绍的工作内容，利用大语言模型的模拟能力，弥合虚拟环境和实际环境的差距。并且，Agent可以根据实时的奖励进行自更新。
* **郭译凡：**从决策的角度对比，LLM是基于可解释的意图生成，RL算法只是基于黑盒的数值映射。在角色支持方面，LLM可以作为一个动态的角色适应多种场景，但RL只能面向给定的场景。LLM还具有很强的泛化能力，可以通过提示词动态快速调整扮演的角色。但相较于RL的计算需求，LLM的推理开销大的多，很难在边缘设备上部署，并且实时响应的时间相对较长。
* **承楠：**要理解大语言模型的Agent要首先理解大语言模型Agent的工作机制，LLM输入和输出的文字有一个作用窗口，在下一次回答问题时可以把上次输出的内容再作为一个输入反馈到下一次的输出中，下一次发生的事情可以同时反应行为和输入的结果，这使得LLM具有一定的记忆能力和学习能力。除了短期的记忆，还具有长期的记忆能力，它是通过某种方式把现实世界中的一些知识给定大模型一些前期的知识贮备，给予LLM一定的行为准则。LLM可以根据自己的任务自己去决定调用什么tool，例如，网络上调用浏览器搜索结果，现实世界中调用一个操纵机械接口。如果想让LLM扮演什么角色，可以给定一些系统的提示词，给定一个长期的令LLM场景遵循的规则。通过调整提示词，就能够避免复杂的模型训练。这也是为什么大模型具有泛化能力的一种原因。角色支持也是建立在这种泛化能力的基础上，
* **郭译凡：**是的，LLM通过记忆能力和调用tools，就能大大提高了LLM的准确性。
* **郭译凡：**MCP架构能够使大模型具有高可组合性质，使Agent之间拥有可交互的能力，能够再不同的平台上操纵不同的Agent。
* **尹志胜：**你这几个性能图为什么大语言模型性能更好。
* **郭译凡：**类似于chatgbt，这种LLM是具有逻辑思考的能力，因此相对于般的强化学习更有逻辑，能够通过结果反作用与回程/。首先，LLM
* **尹志胜：**你这个解决的肯定是没有理论极限，如果是一个有理论极限的结果，那肯定是存在一个对应关系，但是你现在这个LLM对应关系的是如何突破这个指定关系的？你是怎么理解这个次优解的关系？靠什么突破这个次优的改善关系的？
* **郭译凡：**首先，大语言具有一个相对好的初始状态，强化学习从一个完全没有训练的角度出发，因此相对来说具有一些更好的初始点。没有经过预训练就具有一定的优化能力。
* **尹志胜：**你这个预训练是什么意思？
* **郭译凡：**从网上下载下来一些数据直接应用。
* **尹志胜：**但这个数据本身还是有局限性的，无法说你这个解就是好的解，数据本身是杂乱无章的，你的逻辑是什么？我该怎么理解大语言模型的本质。你这个表达的意思就是找问题建模，然后通过问题建模找算法，这不是还是一样的意思吗？这个本质究竟是什么？
* **承楠：**这个大模型是怎么知道这个reward怎么计算的？
* **郭译凡：**这个其实不是独立让大模型去执行这个任务，这里其实这个reward只是给出一个评价标准，让大模型本身去调整这个策略，大模型本身是不改变参数，RL需要根据目标函数一直调整这个参数。
* **周淙浩：**这行加载大模型参数，跟整个算法之间有什么关系，除了一开始提取过参数，后面和大模型完全没关系了。至少从这个工作上来看，很难看出来LLM有什么用。
* **尹志胜：**如果是目前这个优化流程，任何问题都是一股脑的塞进去，然后相当于也不知道整个实现过程。哪怕是大模型输出一些参数，这个参数比一般的好，但是这个是不是真的能实现，怎么确保他能够实现。这个逻辑到底是怎么应用于现实生活的逻辑？例如，强化学习是达到一个次优，那你怎么通过一定范围的解决方式，证明你能够超过一个次优解。如何能够判断这个解决的办法的稳定性呢？他的能力边界，应用的代价又是什么呢？这个逻辑要好好思考。
* **承楠：**未来的网络怎么怎么样，肯定是不适合完全利用LLM替换的，很多的环境可以尝试用LLM去管理和优化，这个肯定是点对点的去证明的。另外一个，现在的LLM和Agent直接放到网络中肯定也不是特别合适，将来肯定还有进一步的发展
* **尹志胜：**现在的很多LLM很多的应用都是一种工程问题，就是把人家训练好的参数拿下来，然后用LLM改善自己的参数，确实也许可行，但是这个问题没法定义，没法去确定自己的适用范围。
* **承楠：**我见过一种应用，用LLM设定一个模拟的社区，社区中有各种模拟的人，这里的模拟的人可能会结婚，可能会住到一起，可能会被辞退，并且具有相应的情绪，这些行为都会变成其他人的输入，同时可以定定义我们输入的环境，这些人还有产生新的行为，怎么依据这些行为去产生新的行为，这就是大语言模型的意义。我认为未来的大模型应该是面向某种类似于需求的任务。
* **王秀程：**这里其实还涉及到一个问题，LLM全都是一种概率模型，他很难对应到那种创建出来的具有唯一解的答案。
* **承楠：**但是我们做的很多那种按需服务其实没有那么强的唯一性，所以可能还是比较适合去做的。
* **周淙浩：**LLM和强化学习这两种方法本身就是两种描述问题的方式，这种环境是很难在一个问题下去找到合适的比较方式的，虽然是描述同一个问题，但是描述的方式不一样。
* **王秀程：**有些问题描述的时候是没法用一个确认的公式去描述的，因此还是很难用一个确认的框架去描述LLM描述需求。
* **周淙浩：**这种要实现其实是可以实现的，但是复杂度太高了。看你怎么理解，大模型构成的一切都可以视为Agent，环境里的每个变量也可以是为Agent，作为LLM的构成部分。
* **承楠：**所以我们还是得深挖LLM得技术细节，去寻找大模型使用的场景，做出对世界的合理解释。
* **周淙浩：**我们也可以学习LLM里建模的方式，基于Agent的方式去描述一个系统，但我们最好还是把Agent和LLM分开来考虑，比如Agent之间的交互以前是基于博弈论来描述的，但是人和人之间的交互是很难通过简单的纳什均衡来描述完的。类似于用神经网络来拟合一个高维的函数，来建模这些Agent之间的关系。
* **尹志胜：**最终来讲的话，这些事情还是有物理映射的，很难突破这种有界问题，比如说我建模成一个模型，我总要用这个模型做一件有意义的事情，但这个事情肯定要依据物理模型的事情，因此这个事情是有界的，干不成很多事情。
* **承楠：**但我们可以通过多个有界的问题协调去解决这个问题。这就需要多思考，多钻研。

**二.会议总结：**

**1.内容**

1).第一部分介绍了传统的RL和LLM的区别，以及LLm的优势

2).介绍了现有的主流LLm的框架，MCP和AA

3).第三部分介绍了一些现有的LLM应用的工作

4).第四部分探讨了一些实际的LLM应用于通信系统的一些挑战和瓶颈。

2.老师们今天提出了关于LLM应用范围的观点和看法，还有质疑都好好理解一下，虽然现有的能力边界和应用范围还不是很清晰，但是未来，我们还是需要一个发展规划，以期将来能够达到这个目标。