Paper Reading 8

沙之洲 2020012408

Self-Refine:

1 这篇文章用产生回答的 LM 作为回答的 evaluator。我们该如何证明 LM 对于回答的 evaluator 是一个合理的 evaluate 呢？有没有可能 refined answer 被 LM 产生的 feedback 错误引导？

2 事实上，经过若干次 LM 的自我修正之后，refined answer 仍然可能是错误的，这可能是因为 LM 本身能力的不足。我们应该如何推断 refined answer 仍然是错误的这个现象，是因为模型 refined 次数不够？还是 LM 本身不具备回答这个问题的能力？换句话说，我们是否可以通过某种方式合理预测经过 LM refined answer 最终能够达到什么样的准确度？

Generative Agent:

1 在 memory retrieval 的过程中，本篇工作采用的是传统的检索方式，根据 state 选择一个 memory stream 的 subset，但是这种方法是不 efficient 的。在很多检索任务上，RL agent 被引入来优化检索。是否可以考虑将 memory retrieval 这个环节用 RL agent 替代，reward 可以设计为 LM 在所选取 subset 上的表现。

2 本篇工作很大篇幅都在处理 memory 的问题。不久之前有一篇 lora 的工作，通过在 LM 外挂规模较小的可训练模块，极大程度提升了 LM finetune 的速度。这里是否可以考虑将 generative agent 的 memory 反向传播回 LM，用外挂的 lora 来存储额外的 memory。这实际上也比较符合人类现有的状态。我们的知识和语言能力是大脑中固定存储的东西，每天的事务会不断地更新。Lora 是否能更好地模拟这种状态？