

CAUSAL Lab 硕士生与博士生申请选拔

张智恒课题组 CAUSAL Lab, SUFE

版本：2026 年 5 月 19 日

题目目的

本题不是考察同学是否已经熟悉因果推断的全部知识，而是考察以下能力：

1. 能否从一个较大的研究方向中提出**清晰、具体、可验证**的科学问题；
2. 能否识别已有方法的边界，并提出具有差异化的改进思路；
3. 能否进行基本的数学建模、实验设计和学术写作；
4. 能否规范、透明、负责任地使用大模型辅助科研，而不是用大模型替代独立判断。

1 背景与大方向

CAUSAL Lab 关注有理论基础并能落地应用的因果学习方法，尤其强调在现实约束下刻画**模型假设、观测数据、可识别边界、统计效率与决策目标**之间的关系。实验室当前重点方向包括（见我的主页）：

- (i) 网络结构下的在线实验设计与推断；
- (ii) 因果推断中的最优传输与几何结构；
- (iii) 面向工业现实约束的因果推断方法论，例如 RCT & OBS 融合、预测后优化、结构化数据类型、LLM 行为建模等。

选拔大方向：

面向真实约束的因果学习：

在网络干扰、有限随机实验、观测偏差、未观测混杂、复杂输出或决策反馈存在时，如何构造**可识别、可估计、可验证、可用于决策**的因果方法？

候选人需围绕上述大方向，选择一个具体切入点，完成一份**小型科研 proposal**。本题允许使用 ChatGPT、Claude、Gemini、Kimi、DeepSeek 等大模型，但必须提交完整 prompt 记录，详见第 4 节。

2 可选切入点

候选人可以从以下切入点中选择一个，也可以自行提出与大方向高度相关的切入点。若自行提出，需要在开头说明其与 CAUSAL Lab 研究方向的对应关系。

方向 A：网络干扰与未知拓扑下的因果推断

现实中的 treatment 往往会通过社交网络、推荐系统、交通网络或市场结构产生溢出效应。然而，真实网络可能未知、动态变化或只能部分观测。请提出一个问题：当网络拓扑未知或不可靠时，如何进行因果效应估计、实验设计或 off-policy evaluation？

可思考的问题包括但不限于：

- 如果真实网络不可观测，但可以观测到行为相似性、地理邻近性或交互日志，如何构造可验证的 exposure mapping？
- 如何在估计精度与实验 regret 之间取得平衡？
- 当 exposure mapping 错设时，如何进行敏感性分析或稳健推断？
- 是否可以设计一个 “unknown topology” 的 simulation 或真实数据 benchmark？

方向 B：有限 RCT 与观测数据融合下的主动采样

工业场景中 RCT 代价高昂，但观测数据充足。请提出一个方法，用尽可能少的随机实验样本，达到接近全量 RCT 的因果估计或策略学习性能。

可思考的问题包括但不限于：

- 如何设计一个可调参数，使 RCT 样本数与估计性能之间形成更优的 trade-off 曲线？
- 在有限 RCT + 大量 OBS 的设定下，如何利用 overlap、domain discrepancy、propensity、proxy outcome 或 representation uncertainty 指导 post-selection？
- 如何证明或实验展示你的方法优于 random subsampling、uncertainty sampling、propensity-based sampling 等 baseline？

方向 C：未观测混杂、partial identification 与最优传输

当 unconfoundedness 或 overlap 不成立时，点估计可能不可识别。请提出一个使用 partial identification、proxy/negative control 或 optimal transport 思想的因果效应界限方法。

可思考的问题包括但不限于：

- 如何把某类不可识别问题转化为一个可计算的优化问题？
- 如何利用协变量、代理变量或边际分布信息收紧 bounds？
- 如何在高维变量或结构化输出下避免 bounds 过宽？
- 能否提出一个简单 theorem，说明某个额外信息会单调地收紧 identification region？

方向 D：分布式结果、风险敏感目标与 policy learning

许多决策不只关心平均收益，还关心结果分布、风险、尾部表现或群体公平。请提出一个面向 distributional outcomes 的 causal policy learning 问题。

可思考的问题包括但不限于：

- 如果 outcome 是一个分布而不是标量，policy value 应如何定义？

- Wasserstein 距离、quantile 函数或 barycenter 是否能给出合适的 welfare?
- 主动采样为什么可能比在标量 outcome 中更自然地适配 distributional outcome?
- 如何设计 regret、estimation error 或 decision instability 的评价指标?

方向 E: LLM/多智能体系统中的因果评估与策略干预

大模型和多智能体系统正在进入推荐、教育、金融、交通和科研辅助场景。请提出一个因果问题，用于评估或干预 LLM/agent 的行为。

可思考的问题包括但不限于：

- 如何定义一个 LLM 行为的 treatment、outcome、confounder 与 interference?
- 当用户、模型与平台策略互相反馈时，传统 causal estimand 如何失效?
- 如何设计小规模实验或离线评估框架，判断某种 prompt、工具调用或推荐策略是否真正提升长期结果?
- 如何避免把相关性误当作因果性?
- 如何探索模型从数据中恢复出真实科学规律的过程? 如何界定模型的唯一识别性?

3 提交内容

请提交一个 PDF 文件，正文篇幅无要求，此外可提交参考文献和附录。可以使用中文或英文写作；若中英文混合，请保持术语统一。注意，因为我们是允许使用大模型辅助作答的，所以垒砌篇幅是无成本行为，不会带来加分。

- R1. 标题与摘要。**摘要不超过 250 字，说明研究问题、核心方法、预期贡献和验证方式。
- R2. 问题定义。**用自然语言解释实际背景，并给出一个基本数学形式。例如，明确 data、treatment、outcome、policy、estimand、assumption、source of bias 中的关键元素。
- R3. 已有工作的边界。**至少阅读 3 篇相关论文，其中至少 1 篇应与 CAUSAL Lab 现有研究方向直接相关。请不要堆砌文献；重点说明已有方法**不能解决什么**。
- R4. 你的创新点。**给出至少两个可能 idea，并选择其中一个展开。需要说明：为什么这个 idea 不是已有方法的简单组合？它最可能失败在哪里？
- R5. 方法框架。**给出算法流程、优化目标、估计器、置信区间、bound 或 policy learning objective 中至少一项。鼓励使用清晰的伪代码或结构图。
- R6. 理论或半理论分析。**不要求完整证明，但必须写出一个形式化命题、lemma、proposition 或 conjecture，并给出证明思路。命题可以很小，但必须准确。
- R7. 实验或验证方案。**说明数据来源、simulation 机制、baseline、评价指标和 ablation。若暂时无法做实验，也要说明如何验证该 idea 是错的或有效的。
- R8. 半年执行计划。**写出如果加入课题组，半年内如何推进该题目，包括阅读、复现、理论推导、实验和写作计划。

R9. 大模型使用记录。按第 4 节提交完整 prompt log。

强制要求

你的 proposal 必须包含以下四个元素：

- (1) 一个明确的 causal estimand 或 policy value；
- (2) 一个你认为最关键、也最可能被违反的假设；
- (3) 一个可以区分你方法与 baseline 的实验或反例；
- (4) 一段 “Why this may fail”，诚实说明你的方法可能无效的情形。

缺少任意一项，将被视为不完整提交。

4 大模型使用规范

本题允许使用大模型。我们更关心你如何提问、如何验证、如何选择，而不是是否使用大模型。请在附录中提交完整 prompt 记录，格式如下。

时间	模型/工具	完整 Prompt	如何使用输出
2026-xx-xx	ChatGPT Claude / etc.	/ 请逐字粘贴你的 prompt, 不要只写“润色文章”或“帮我想 idea”。	例如：用于生成候选文献清单；我人工删除了不可靠文献，并重新检查了定理表述。

表 1: 大模型使用记录模板。可以增加行数。若未使用大模型，也请明确写“未使用”。

允许的使用方式：

- 用于解释背景知识、生成初始文献关键词、润色语言、检查逻辑漏洞；
- 用于生成代码草稿，但你必须理解并能解释代码；
- 用于反驳自己的 idea，例如要求模型扮演审稿人指出缺陷。

不允许的使用方式：

- 编造不存在的论文、定理、实验结果或引用；
- 不经核查地复制大模型生成的数学证明；
- 只提交一篇通用综述，没有自己的问题定义和判断；
- 隐瞒关键 prompt，或只写“使用大模型辅助完成”。

5 提交格式

请提交一个压缩包，命名为

Name_University_CAUSALMasterProposal.zip.

压缩包中应包含：

1. `proposal.pdf`：正文；
2. `source.tex`：LaTeX 源文件；
3. `prompt_log.pdf` 或 `prompt_log.md`：完整大模型 prompt 记录；
4. 可选：`code/` 文件夹，包含 `simulation` 或初步实验代码；
5. 可选：`figures/` 文件夹，包含图表源文件。

最后提醒

一个好的提交不一定需要提出宏大的理论。更重要的是：问题要具体，假设要清楚，方法要被验证，失败情形要诚实。我们希望看到候选人如何思考，而不仅是看到一个看似完整但无法落地的答案。