

2025 Lean 与数学形式化讲义(1A)

上海交通大学 AI4MATH 团队

0 相聚于东海之滨

0.1 具体课程安排

	7/14 (一)	7/15 (二)	7/16 (三)	7/17 (四)	7/18 (五)	7/19 (六)
9:30 – 11:30	课程 1A	校园参观	课程 3A	课程 4A	课题实践	课题展示
11:30 – 15:00	午休	午休	午休	午休	午休	午休
15:00 – 17:00	课程 1B	课程 2B	课程 3B	课程 4B	课题实践	课题展示
17:00 – 18:00	晚休	晚休	晚休	晚休	晚休	晚休
18:00 – 20:00	课程 1C	课程 2C	课程 3C	课程 4C	课题实践	机动

* 除校园参观外, 其他活动均在光彪楼 206 教室进行.

* 午休时间和晚上 20:00 – 21:00 可继续留在教室进行课题实践或答疑.

课程 1A: 相聚于东海之滨

主讲: 刘晓洋

课程材料: 讲义 1A, 生活手册

主要内容:

- 1. 分发名牌;
- 2. 介绍具体课程安排与分组课题内容;
- 3. 形式化相关数学史介绍;
- 4. AI4Math 领域研究现状.

课程 1B: 初识 Lean

主讲: 刘云天 (猫猫)

课程材料: 讲义 1B, MIL C02

主要内容:

- 1. 初识类型论;
- 2. 使用 Lean 进行形式化数学证明.

课程 1C: 一阶形式逻辑

主讲: 刘云天 (猫猫)

课程材料: 讲义 1C, MIL C03, 1C.lean

主要内容:

- 1. 使用 tactic 处理一阶形式逻辑命题;
- 2. 自然语言证明的形式化.

课程 2A: 校园参观

课程材料: 生活手册

主要内容:

1. 参观校史馆;
2. 课题分组与组内破冰活动;
3. 幸运玩家可以提前试玩到猫猫精心设计的 Lean 桌游.

课程 2B: 朴素集合论

主讲: 刘云天 (猫猫)

课程材料: 讲义 2B, MIL C04, Mathlib Explorer

主要内容:

1. Mathlib 简介与定理查找方法;
2. `Mathlib.Data.Set` 中朴素集合论的基本定义和定理.

课程 2C: 练习答疑

课程材料: 讲义 1C + 2B, MIL C03 + C04

主要内容:

1. 完成 MIL C03 和 C04 两节的半数练习题;
2. 答疑.

课程 3A: 依值类型论

主讲: 朱韬

课程材料: 讲义 3A, `3A.lean`

主要内容:

1. 依值类型论;
2. Lean 中项的构造语法;
3. Curry-Howard 同构.

课程 3B: 策略证明原理

主讲: 朱韬

课程材料: 讲义 3B, `3B.lean`

主要内容:

1. 依值类型论视角下的基本策略介绍;
2. 策略的进阶用法.

课程 3C: 练习答疑

课程材料: 之前各项讲义, MIL C03 + C04

主要内容:

1. 完成 MIL C03 和 C04 两节练习题;
2. 答疑.

课程 4A: 初等数论

主讲: 王浩丞

课程材料: 讲义 4A, MIL C05, 4A.lean

主要内容: Mathlib 中数论基础概念、核心定理、同余模运算、数论函数及经典问题.

课程 4B: 抽象代数

主讲: 王浩丞

课程材料: 讲义 4B, MIL C09, 4B.lean

主要内容: Mathlib 中基础代数结构 (群, 环, 模) 的定义及相关应用.

课程 4C: 数学分析

主讲: 白一凡

课程材料: 讲义 4C, MIL C12 + C13, 4C.lean

主要内容: Mathlib 中极限定义, 度量空间与赋范空间, 连续线性算子, 简单微积分.

0.2 课题内容

我们提供两种课题形式: 分组课题和个人课题. 个人课题是作为分组课题的补充引入的, 除非真的特别社恐或者你真的对 Lean 已经了如指掌, 还是推荐大家参加分组课题.

需要注意, 有些工作可能是前人从未完成过的, 或是前人做过却由于种种原因无人问津的事情. 在课题开展过程中可能会遇到各种各样难以预料的困难. 由于活动时间有限, 制定计划的时候一定先确定一个容易完成的核心目标, 例如证明定理的一些特例, 或是适当牺牲理论框架的泛用性等, 在有时间的情况下再考虑扩展内容. 相信与形式化领域有缘的人, 终会找到属于自己的方向并有所成就.

0.2.1 分组课题

课题 1 竞赛题证明形式化

请选择一道高中竞赛及以上复杂度的数学题目, 将题干形式化后, 完成相应的形式化证明.

以下建议供参考:

1. 选择题目的时候, 尽量选择涉及理论框架比较简单, 比较容易形式化描述的数论或不等式题目. 一些纯几何内容可能在 Mathlib 中没有对应的库, 而一些组合数学问题可能难以形式化, 这些会给形式化带来困难. 一些复杂的求和算子也可能难以处理.
2. 建议先用自然语言完成证明, 或者选择已有现成自然语言证明的题目, 在此基础上针对形式化进行表述上的优化, 这样可以省去从零开始解决问题的时间. 找到自然语言证明之后, 对其中涉及的各个算子和相关的关键定理在 Mathlib 库中进行查找, 确认它们已经有了对应的实现.

如果没有, 可以尝试自己实现 (但这会十分麻烦), 或者尝试在 Mathlib 中寻找相近的算子进行替代.

3. 进行形式化时, 可以先进行引理分割, 先使用 `sorry` 占位符勾勒大体的证明框架, 再交由组员分工完成各个部分的证明.

课题 2 数学问题形式化

请选择一个比较前沿的数学问题, 表示出其所涉及的各个数学概念的定义, 找到它们在 Mathlib 中的实现, 或是借助 Mathlib 中已有的概念自己完成定义链条, 勾勒一个定理库的大致框架. 不需要完成完整的形式化证明, 但可以对一些关键的性质或是引理给出证明.

以下建议供参考:

1. 不要选择过于复杂的数学问题. 与前一个课题类似, 如果涉及的理论过于复杂, 可能会导致形式化的困难.
2. 如果你选择形式化 Goldbach 猜想之类表示十分初等的问题, 还是建议附带完整的数学证明.

课题 3 Mathlib 学员

Mathlib 库为了确保实用性, 许多数学概念都力求保证其泛用性. 但这也导致了許多数学概念的处理十分复杂, 甚至有些处理方式与数学学习者通常的直觉相去甚远, 让 Lean 的学习和使用成本都十分高昂, 也就很难让 Lean 服务于一般的数学教学和研究. 但 Lean 原本是一个十分强大的工具, 不应完全被限制在现有 Mathlib 库的框架之内. 选择一段数学理论框架, 例如借助 Mathlib 中已有的实数概念, 用简单的 $\epsilon - \delta$ 语言描述数学分析等学科, 不妨尝试在不引入过于抽象的代数和拓扑工具的情况下形式化极限及函数连续性等简单概念, 让 Lean 融入到本科生的数学学习当中.

课题 4 物理形式化

尝试将一些简单的物理学概念形式化为数学对象. 这会是一个比较有挑战性的课题, 因为从零开始构造对象需要大量运用归纳类型, 需要对待形式化领域的对象有很清晰的理解.

以下建议供参考:

1. Physics in Lean 项目是一个正在进行的物理学形式化项目, 也许可以提供一些灵感.
2. 不妨从简单的概念框架开始, 例如以简单的牛顿力学作一个初步的尝试.

课题 5 其他任何你想借助形式化工具做的小组课题!

五子棋, 围棋, 红石电路, 三国杀... 都可以, 不过一定把握好复杂度, 不要整太难了做不出来.

0.2.2 个人课题

课题 6 Mathlib 文档整理

选择一个 Mathlib 文件, 用自然语言阐释各个条目的内容, 整理成 NL-FL 平行的 LaTeX 文档.

课题 7 (专题二) 数值代数理论框架蓝图

为专题二数值代数学科的形式化做准备.