## 第一章：导论

我们为什么需要新的道德理论？

这是因为，尽管世界上现存的各种道德理论在日常生活中的许多判断结果看似相似，但它们背后的逻辑基础却大相径庭。每一套体系都源自不同的哲学假设、价值观或对人性的理解方式，而这些差异在遇到极端伦理情境时会暴露出它们各自的局限，导致它们在面对某些关键道德难题时变得自相矛盾或难以令人信服。

这些道德悖论正是传统伦理理论难以解释或自洽的“症状”。它们不仅暴露了各体系在边界处的逻辑崩溃，也让人怀疑：我们是否真的理解了“道德”的本质。

最著名的例子莫过于电车难题（Trolley Problem）。问题设置是这样的：一辆失控的电车即将撞死前方轨道上的五个人，而你可以拉动一根拉杆，将电车转向另一条轨道，但这将导致另一名无辜的人被撞死。你该不该拉杆？

这个问题之所以被称为“难题”，正是因为各大伦理理论都对它给出了不同的、彼此冲突、却又都显得不完全令人满意的答案：

**各伦理流派的回答及其局限性：**

功利主义：主张拉杆，因为牺牲一人救五人“总体幸福最大化”。

问题：这种逻辑若推广，会导致容许杀死一个健康人来移植器官给五人的荒谬结论。它忽视了社会信任和个体尊严的问题。

义务论（康德主义）：主张不可拉杆，因为主动杀死无辜者违反了道德义务。

问题：这一立场即使在“牺牲一人可救亿万人”的情境下也拒绝让步，显得僵化、脱离现实、反直觉。

双重效应原则：拉杆是可接受的（因为死亡是副作用），但推人下桥是不可接受的（因为死亡是直接手段）。

问题：行为后果与动机之间的区分过于模糊，难以为现代政策或算法提供清晰判断标准。

美德伦理：强调道德行为背后的动机，比如怜悯或勇敢，但回避行为结果的评估。

问题：缺乏判断机制，在实践中难以形成系统的规范制定工具。

社会契约论：尝试从群体共识角度出发判断行为正当性。

问题：在极端情境下（如极少数人做出决定时），“共识”难以衡量，依然缺乏理论一致性。

综上所述，我们发现：

每一种传统道德理论都在电车难题中暴露出内部逻辑的不完整性。

它们不是没有“道理”，但没有一种能在所有情况中保持一致性、可操作性、与社会稳定性的兼容。而道德体系，若要指导一个高度复杂、技术密集、全球化的现代社会，仅仅依靠理想主义或抽象原则是远远不够的。

因此，我们需要一个新的道德理论：

一个基于现实人类心理与社会演化机制之上的理论；

一个在逻辑上封闭自洽、在实践中可预测可调整的系统；

一个可以解释道德悖论而不依赖变通规则的结构；

一个可以用作立法、算法判断、AI伦理决策的通用框架。

这本书将提出这样一套体系，将会重新定义道德，用一种逻辑上面完备，不会出现各种局限性的道德体系。

这将是一个完全不同的出发点，也将带来不同的答案。