开题感想记录

德州扑克是一项标志情商、智慧和勇气的牌类游戏，自20世纪初诞生便风靡全球，深受各个不同职业群体玩家的青睐。由于德州扑克非完备信息特性，游戏玩家都有未知的场上信息，同时还有诈唬等心理战术。这些特点使得德州扑克更贴近生活中真实的拍卖、谈判等金融政治博弈场景，所以德州扑克也许是测试人工智能是否能够应付这些情景的一个平台。

前几年，AlphaGo兴起了机器学习的热潮。在围棋这种完全信息的零和博弈中，作为算法需要解决的仅仅只是如何搜索大规模博弈树的问题，但是在德州扑克这种非完备信息博弈的问题中往往还藏有其他手段。解决此类非完备信息博弈，如果解决的思路是从寻找纳什均衡出发，那么解决手段的基础就是CFR(虚拟遗憾最小化算法)。虚拟遗憾最⼩化算法是⽬前最具有竞争性应⽤最⼴泛的算法。CFR 是⼀种通过两种遗憾最⼩化算法之间反复训练的过程来逼近扩展式博弈纳什均衡的迭代⽅法。通过多次迭代，将整体遗憾分解到各个独⽴ 的信息集中计算局部最⼩后悔值，后悔值是算法没有选择最佳策略⽽损失的游戏收益，只有在执⾏策略后才知道。遗憾最⼩化算法可以使后悔值随着游戏进⾏呈线性增长，最终实现与采⽤最优策略相同的效果。CFR的关键点在于不会存储呈指数增长的决策数量的后悔值，⽽是存储和最⼩化在每个信息集和相应后续决策中的不断修改调整的后悔值，这样就可得出任何决策下的后悔值。通过在每⼀次迭代中对每个玩家的策略取平均值，可以近似求出纳什均衡，并且随着迭代次数的增加，近似程度也会提⾼。当前在两人有限注德州扑克的问题上利用CFR算法已取得较高成就，通过本次课题，我能够了解德州扑克的规则，掌握德州扑克的一般玩法；理解CFR算法思路，并且能够编写CFR的主要代码；同时更加熟悉Python语言编程，熟悉面向对象编程，构思好主要框架，写出整洁美观实用性强可移植性强bug少的代码。对我而言是一次不错的挑战。两人有限注德州扑克取得的突破性进展也鼓舞了无数学者，研究 CFR 算法在非完备信息机器博弈中的应用具有广泛的理论研究价值和深远的实际应用意义。