读书笔记 - 人工智能：现代方法

第二部分：问题求解

第 3 章：通过搜索解决问题

问题求解智能体：智能体通过**搜索**形成通往目标状态路径的动作序列。

模型概述：使用原子表示，本章只考虑简单的环境：回合式的、单智能体的、完全可观测的、确定性的、静态的、离散的和已知的环境；问题的解是一个动作序列

智能体的求解过程

1. 目标形式化
2. 问题形式化
3. 搜索
4. 执行

搜索问题的形式化定义

- 模型：
  - 状态空间
  - 初始状态
  - 目标状态
  - 智能体可以采取的动作
  - 转移模型
  - 动作代价函数
- 一个动作序列形成一条路径 (path)，而解 (solution) 是一条从初始状态到某个目标状态的路径。

搜索算法分类

- **无信息搜索（盲目搜索）**：仅依赖问题定义，无额外启发信息。
  - **广度优先搜索（BFS）**：完备且最优（单位成本），但空间复杂度高（$O(b^d)$）。
  - **一致成本搜索（UCS）**：按路径成本 $g(n)$ 扩展，适用于非单位成本。
  - **深度优先搜索（DFS）**：空间效率高（$O(bm)$），但不完备（可能陷入无限路径）。
  - **迭代加深搜索（IDS）**：结合 BFS 的完备性与 DFS 的空间效率，时间复杂度与 BFS 相当。
  - **双向搜索**：从初始状态和目标状态同时搜索，减少扩展节点数。
- **有信息搜索（启发式搜索）**：利用启发函数 $h(n)$ 估计到目标的代价。
  - **贪婪最佳优先搜索**：优先扩展 $h(n)$ 最小的节点，可能非最优。
  - **A* 搜索**：最小化** $f(n) = g(n) + h(n)$**，若** $h(n)$ 可采纳（不超估）**且**一致 **，则 A* 完备且最优。
  - **IDA**：迭代加深的 A，节省内存。
  - **加权 A***：通过调整启发式权重（$f(n) = g(n) + W \cdot h(n)$）加速搜索，但牺牲最优性。

启发式函数设计

- **可采纳性**：$h(n)$ 不超估真实代价。
- **一致性**：满足三角不等式 $h(n) \le c(n, a, n') + h(n')$。
- **生成方法**：
  - **松弛问题**：通过简化规则（如忽略移动限制）获得可采纳启发式。
  - **模式数据库**：预计算子问题的解代价（如 8-puzzle 的局部布局）。
  - **地标点（Landmarks）**：预计算关键节点间的最短路径，用于快速估计。

性能评估

- **完备性**：能否找到解（若存在）。
- **最优性**：能否找到最低成本解。
- **时间与空间复杂度**：通常以分支因子 b 和解深度 d 表示。
  - **有效分支因子** b^*：衡量启发式质量的指标。

第 4 章　复杂环境中的搜索

模型概述：放松第三章的限制。

局部搜索：

寻找最终状态而不考虑到达该状态的路径；相关算法：

- **爬山算法**：基本的局部搜索技术，每次迭代移动到具有最高值的相邻状态，直至达到局部最大值。但它容易陷入局部最优、山脊和高原等困境。为解决这些问题，出现了随机爬山、首次选择爬山和随机重启爬山等变体算法。
- **模拟退火算法**：结合爬山算法和随机漫步，通过控制温度参数接受“下坡”移动，能以一定概率跳出局部最优，最终找到全局最优解，在 VLSI 布局、工厂调度等领域应用广泛。
- **局部束搜索**：跟踪 k 个状态，从 k 个随机生成的状态开始，每次迭代生成所有 k 个状态的后继状态，选择 k 个最佳后继状态继续搜索。随机束搜索是其变体，通过按比例选择后继状态增加多样性。
- **进化算法**：受生物自然选择启发的随机束搜索变体，通过种群个体的选择、重组和变异来寻找最优解。遗传算法是其中一种，个体以字符串表示，通过交叉和变异操作生成新个体，在复杂结构化问题优化中应用较广。

处理连续空间

- 现实中很多环境是连续的，连续动作空间分支因子无限，多数传统算法难以处理。
- 可通过离散化连续空间或使用经验梯度方法来解决，还可利用微积分计算梯度寻找最大值。牛顿 - 拉夫森方法是常用的求解函数根的技术，在寻找函数极值时有广泛应用。此外，还介绍了约束优化问题，线性规划和凸优化是其中重要的类别。

非确定性动作的搜索

- **不稳定吸尘器世界**：引入非确定性后，吸尘器世界的解决方案不再是简单的动作序列，而是条件规划。例如在不稳定吸尘器世界中，需要根据不同的状态执行不同的动作。
- **AND-OR 搜索树**：用于解决非确定性问题，树中包含 OR 节点（由智能体的动作选择产生分支）和 AND 节点（由环境的不确定性产生分支）。解决方案是满足一定条件的子树，通过深度优先等算法搜索 AND - OR 图来找到解决方案。
- **循环规划**：在某些非确定性环境中，如滑溜的吸尘器世界，可能需要循环规划来解决问题。但循环规划是否有效取决于非确定性的原因，如果是由于随机因素导致的，重复动作可能会成功；如果是由于未观察到的事实导致的，重复动作可能无济于事。

部分可观测环境中的搜索

- **无观测搜索**：即传感器缺失问题，虽然智能体没有感知信息，但仍可通过在信念状态空间搜索来解决问题。可将物理问题转化为信念状态问题，使用普通搜索算法求解，还可通过剪枝等方法提高效率。此外，还有增量信念状态搜索算法，能更快速地检测问题是否可解。
- **部分可观测环境搜索**：部分可观察环境中，智能体通过感知获取部分信息，信念状态会根据动作和感知不断更新。通过预测、计算可能感知和更新三个阶段来推导信念状态的转换，使用 AND - OR 搜索算法可得到条件规划的解决方案。
- **部分可观测环境的智能体**：在部分可观察环境中，智能体执行条件规划并维护信念状态。例如在幼儿园吸尘器世界和机器人定位问题中，智能体根据感知和动作更新信念状态，以实现目标。

在线搜索智能体与未知环境

- **离线搜索智能体**：在执行第一个动作之前就已经计算出一个完整的解。
- **在线搜索智能体**：交替进行计算和动作：它首先执行一个动作，然后观测环境并计算下一个动作。
- 在线搜索智能体在执行动作后根据感知更新环境地图，深度优先搜索是常用的在线搜索策略。在线深度优先探索智能体通过记录动作结果和回溯来探索环境，但只能在动作可逆的环境中工作。
- **在线局部搜索**：爬山搜索是一种在线局部搜索算法，但容易陷入局部最大值。随机漫步可用于探索环境，但效率较低。学习实时 A$*$（LRTA$*$）算法通过存储成本估计值并根据经验更新，能更有效地探索环境，在有限、安全可探索的环境中可保证找到目标。
- **在线搜索中的学习**：在线搜索智能体可通过记录经验学习环境“地图”，并通过局部更新规则获取更准确的成本估计。此外，还可通过增量搜索等方式为未来的搜索问题做准备，提高搜索效率。

第 5 章：约束满足问题（CSP）

模型概述

- **约束满足问题（CSP）**：通过**一组变量（即状态的因子表示）**及其对应的域和约束条件来描述问题。目标是找到满足所有约束的变量**赋值**。
- **组成部分**：
  - 变量集合 $ X $。
  - 域集合 $ D $，每个变量对应一个允许的取值集合。
  - 约束集合 $ C $，定义变量之间的合法组合。

示例问题

- **地图着色问题**：为地图区域着色，相邻区域颜色不同。
- **作业车间调度问题**：安排任务顺序，满足时间约束和资源限制。
- **数独问题**：填充数字，满足行、列和子格内的唯一性约束。

约束传播与局部一致性

- **节点一致性**：变量的域值满足一元约束。
- **弧一致性**：变量的每个值在其邻居变量的域中存在相容值。
- **路径一致性**：扩展至三元组变量的一致性检查。
- **k-一致性**：推广至更高维度的约束一致性。
- **全局约束**：如 `Alldiff` 和资源约束，用于高效处理复杂约束。

CSP 的搜索

有时我们完成约束传播过程后仍存在具有多个可能值的变量。在这种情况下，我们必须通过搜索来求解问题。

回溯搜索

- **基本思想**：通过深度优先搜索尝试变量赋值，遇到冲突时回溯。
- **启发式**：
  - **最小剩余值（MRV）**：选择剩余合法值最少的变量。
  - **度启发式**：选择约束最多的变量。
  - **最少约束值**：选择对邻居变量限制最少的值。
- **智能回溯**：
  - **冲突导向回溯**：根据冲突集直接回溯到问题源头。
  - **约束学习**：记录冲突以避免重复搜索。

局部搜索

- **最小冲突启发式**：选择使冲突数最小的赋值，适用于大规模问题。
- **优势**：内存占用低，适合动态变化的 CSP。

问题结构

图结构

- **独立子问题**：分解为互不影响的子问题，分别求解。
- **树结构 CSP**：无环约束图可通过拓扑排序高效求解。
- **割集条件**：通过移除部分变量将图转化为树结构。
- **树分解**：将约束图分解为树状结构，子问题间共享变量。

约束结构：

- **对称性约束**：通过引入额外约束消除对称解，减少搜索空间。

第 6 章：对抗搜索与博弈

模型

竞争环境，两个或以上的智能体具有互相冲突的目标，这引出了对抗搜索问题，如棋类游戏。

多智能体环境的建模：1. 把多智能体看作一个经济（economy）整体来考虑，不需要预测单个智能体的动作 2. 认为对抗智能体是环境的一部分，这部分使得环境变成不确定的 3. 用对抗博弈树搜索技术显式地对对抗智能体建模。即本章所涵盖的内容。

博弈论（Game Theory）

- **博弈的分类**：
  - **确定性 vs. 随机性**（如国际象棋 vs. 双陆棋）。
  - **完全信息 vs. 不完全信息**（如围棋 vs. 扑克）。
  - **零和博弈**（一方收益等于另一方损失）。

博弈中的最优决策（Optimal Decisions in Games）

极小化极大算法（Minimax Search）

- **目标**：MAX 玩家选择最大化自身收益的策略，MIN 玩家选择最小化 MAX 收益的策略。

- **局限性**：计算复杂度 $ O(b^m) $，无法处理复杂游戏（如国际象棋）。

Alpha-Beta 剪枝（Alpha-Beta Pruning）

- **优化极小化极大搜索**，剪除不影响最终决策的分支。
- **关键思想**：维护 α （MAX 的下界）和 β（MIN 的上界），提前终止无效搜索。
- **最佳情况**：时间复杂度降至 $ O(b^{m/2}) $。

多人博弈

- 使用**向量评估函数**（如 $[v_A, v_B, v_C]$）代替单一值。
- 联盟可能形成，但需考虑背叛风险。

启发式 Alpha-Beta 树搜索（Heuristic Alpha-Beta Tree Search）

评估函数（Evaluation Functions）

- **替代完整搜索**，在非终止节点估算胜率（如国际象棋的"棋子价值"）。
- **线性加权模型**：

$$\text{EVAL}(s) = w_1 f_1(s) + w_2 f_2(s) + \cdots + w_n f_n(s)$$

  – 例如：兵 =1，马 =3，车 =5，后 =9。

截断搜索（Cutting Off Search）

- **深度限制**：固定深度（如 5 层）或迭代加深（逐步增加深度）。
- **静态 vs. 动态截断**：
  – **静态**：固定深度。
  – **动态**：根据局面稳定性调整（如"静止搜索"避免激烈变化）。

前向剪枝（Forward Pruning）

- **Beam Search**：每层仅保留前 $ n $ 个最佳动作。
- **概率剪枝（PROBCUT）**：基于历史数据预测无效分支。

蒙特卡洛树搜索（Monte Carlo Tree Search, MCTS）

- **适用场景**：高分支因子（如围棋）或难以定义评估函数。
- **核心步骤**：
  1. **选择（Selection）**：从根节点出发，按 UCB1 公式选择子节点。
  2. **扩展（Expansion）**：添加新子节点。
  3. **模拟（Simulation）**：随机对局至终止状态。
  4. **反向传播（Backpropagation）**：更新路径上的统计数据。

随机博弈（Stochastic Games）

- **引入机会节点**（如掷骰子）。

- **期望极小化极大算法（Expectiminimax）**：

$$
\text{EXPECTIMINIMAX}(s) = \begin{cases} \text{UTILITY}(s) & \text{终止状态} \\ \max_a \text{EXPECTIMINIMAX}(\text{RESULT}(s,a)) & \text{MAX 回合} \\ \min_a \text{EXPECTIMINIMAX}(\text{RESULT}(s,a)) & \text{MIN 回合} \\ \sum_r P(r)\text{EXPECTIMINIMAX}(\text{RESULT}(s,r)) & \text{机会节点} \end{cases}
$$

- **示例**：双陆棋（Backgammon）需考虑骰子概率。

部分可观测博弈（Partially Observable Games）

Kriegspiel（不完全信息象棋）

- 玩家仅知己方棋子位置，需推理对手状态。
- **信念状态（Belief State）**：所有可能的合法棋盘状态。
- **策略类型**：
  - **必胜策略（Guaranteed Checkmate）**：无论对手如何应对均能获胜。
  - **概率必胜（Probabilistic Checkmate）**：随机化策略迫使对手犯错。

扑克等卡牌游戏

- **隐藏信息**（如对手手牌）需概率推理。
- **均衡策略**：混合策略（如虚张声势）以迷惑对手。

6.7 游戏搜索算法的局限性（Limitations of Game Search Algorithms）

1. **评估函数误差**：可能导致错误决策。
2. **计算资源限制**：无法覆盖所有可能性（如围棋 $ 10^{170} $ 状态）。
3. **信息不完全性**：部分可观测博弈需复杂推理。
4. **高层策略缺失**：人类直觉（如“控制中心”）难以编码。