拓扑排序

简介

拓扑排序是将偏序的数据线性化的一种排序方法。

复习下偏序和全序的概念:

全序关系是偏序关系的一个子集。

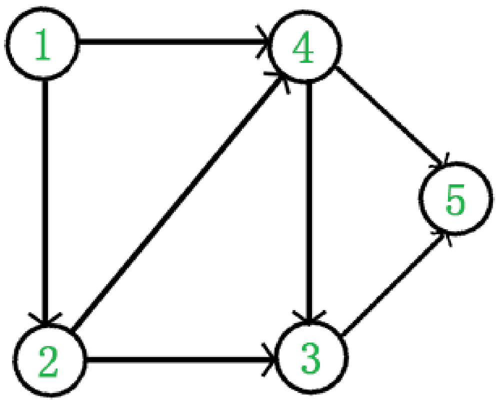
全序是集合内任何一对元素都是可比较的，比如数轴上的点都具有一个线性的数值，因此根据数值就可以进行比较。

偏序是集合内不是所有元素都是可以比较的，比如平面内的点由横坐标和纵坐标组成，是不可直接比较大小的。这是因为横坐标和纵坐标是两个维度，在每个维度内都可以用数值比较，但是维度之间不可量化比较（就像学习成绩和身体素质之间无法量化比较）。当然偏序是个数学概念，未必是多维度引发的不可比较，只需满足以下关系即满足偏序关系:

> 设 P 是集合，P 上的二元关系“≤”满足以下三个条件，则称“≤”是 P 上的偏序关系（或部分序关系）:
>
> （1）自反性：$a \leq a$，$\forall a \in P$;
>
> （2）反对称性：$\forall a$，$b \in P$，若 $a \leq b$ 且 $b \leq a$，则 $a=b$;
>
> （3）传递性：$\forall a$，b，$c \in P$，若 $a \leq b$ 且 $b \leq c$，则 $a \leq c$;

注意这里的“≤”是一个自定义的二元运算符，而不是通常的线性运算的大小关系。

理解了偏序关系之后，拓扑排序就是将偏序关系线性化。举一个具体场景，在有向无环图中，“节点 A 是否可由节点 B 到达”即是一种偏序关系。在该有向无环图中节点，在不移动的情况下可以到达自身，因此满足自反性。且在有向无环图中不存在环，若 A 可到达 B 且 B 可到达 A，则 A，B 必是相同节点，因此满足反对称性。当节点 A 可以到达节点 B，且节点 B 可以到达节点 C 时，节点 A 也可以到达节点 C，因此满足传递性。



在该场景下，拓扑排序即是将有向无环图中所有节点按照“节点 A 是否可由节点 B 到达”来进行线性化排序。如上图所示，对它进行拓扑排序可以使用深度优先算法完成，深度优先算法可以复习课件，其过程大概如下。

1. 选取顶点 s（一般选取入度比较小的节点更合适，比如上图的节点 1），标记状态
2. 若 s 有未被访问的邻居，则选择邻居标记状态继续访问，否则返回

3. 如果一次深度优先搜索还有节点未被访问，则重复步骤 1，直到所有节点被访问到
这种方法可以将满足偏序关系的有向无环图线性化出一种排序结果：1，2，4，3，5。当然对于更复杂的有向无环图可能有多种合法的排序结果。

作业要求

写一个算法可以将有向无环图按照“节点 A 是否可由节点 B 到达”的偏序关系的一种拓扑排序打印出来。示例输入和输出可见 input.txt 和 output.txt（你也可以自己构造用例），你需要首先根据输入文件构建有向无环图，再打印出拓扑排序。