1. 问题重述

1.1 问题背景

设想一种沙漠穿越游戏，其流程大体如下：在遵守八条游戏规则的前提下，玩家开局有一张全局沙漠地图，凭着初始资金购买一定的水和食物。从起点出发，在沙漠中行走，到终点结束。路途中会有三种不同天气，玩家也可以在村庄、矿山补充资源或资金。目标是保证能在规定时间内到达终点，获得尽可能多的资金。

1.2 问题要求

因此本文依据游戏的不同设定，建立数学模型，解决以下问题：

**问题1：**假如仅有一名玩家，并事先知道在整个游戏时间段内的每日天气状况，建立数学模型，试着给出在一般情况下玩家的最优策略。

**问题2：**假如仅有一名玩家，玩家依据仅有的当日天气状况对当天的行动方案进行决策，建立数学模型，试着给出在一般情况下玩家的最优策略。

**问题3：**假如现有n名玩家，游戏规则有一定的变化，据此条件解决以下问题：

**（1）**假如所有玩家事先知道在整个游戏时间段内的每日天气状况，在游戏开始时需确定行动方案且在游戏过程中不能更改。建立数学模型，试着给出在一般情况下玩家应采取的策略。

**（2）**假如所有玩家仅知道当日的天气状况，从第天起，各位玩家在当日行动结束后均知道其他玩家剩余的资源数量以及当日的行动方案，从而确定各自在第二天的行动方案。建立数学模型，试着给出在一般情况下玩家应采取的策略。

二、问题分析

作为游戏玩家，希望在游戏中可以在保障到达终点的情况下尽可能多的挣到钱，因此建立适当的数学模型可以为玩家提供更好的游戏建议与游戏体验。

问题一分析：

问题一要求我们求解在玩家已知所有天气情况下的最优策略。在本游戏中不仅有天数限制，而且还有生存限制，因此在建立模型时应充分考虑这两点。由于天气状况已知，则玩家可以避免到达不了以及生活必需品不够等危险情况。在保证安全前提下，建立路径规划的规划方程组。由于直接计算运算量太大，因此求解本题结果可使用相关智能算法进行求解。

问题二分析：

问题二要求我们在仅知道当天天气情况下，给出玩家的策略。考虑到未来不确定因素较大，因此建议玩家在游戏刚开始尽可能多的购买生活必需品。对于一般玩家可以选择直接去终点或者选择先挖矿后去终点。如果选择了先挖矿后去终点则可以将整个游戏分为两个阶段：起点到矿山，矿山到终点。在起到到矿山过程中玩家只需要决策当天走还是停，因此本阶段可以设立风险度指标针对不同玩家的心理阈值进行决策。在矿山到终点过程中，玩家需要考虑是否挖矿、是否补给、是否去终点等问题，因此可以建立多个决策变量来决策玩家行动。

问题三分析：

问题三要求我们多人游戏中在已知天气情况下的游戏策略。由于多人游戏中改变了游戏规则，则玩家应该尽量避免与其他玩家相遇，从而减少损失。因为天气已知，则可以通过模型一求解出游戏关卡的最优策略，但是为了尽量避免与其他玩家接触，因此本题可以引入静态博弈论模型。在保证所有玩家最终受益均衡的状态下，通过博弈论选出不同的路线，从而最大程度上减少了多个玩家相遇的情况。

问题四分析：

问题四要求我们在只知道当前天气状态下的多人游戏策略。由于未来天气未知，因此玩家可以根据模型二进行每一步的决策，但为了避免玩家相遇情况，本题可以引入动态博弈论进行求解。首先每名玩家都可以根据当前情况去决定第二天的行动，但考虑到存在其他玩家，因此在每一个做决定的过程中都引入动态博弈环节，从而保证了所有玩家收益均衡。

三、模型假设与约定

1）假设在负重上限之下，负重的差异不影响资源消耗的差异；

2）假设在有多名玩家的游戏设定上，所有玩家都会前往矿山；

3）

4）

四、符号说明及名词定义

五、问题一模型建立与求解

5.1 模型准备

问题一要求单个玩家在知道30天每一天的天气下求出最优策略，那么首先应该了解到单人玩家进行游戏有个非常重要的特征：无人争抢最佳路线。那么在研究问题和建立模型可以先将地图抽象化成几个主要点的链接地图，大大减少求解时的寻路空间。如下图以第一关为例，将地图抽象化：

抽象化后几个关键点，起点、村庄、矿山和终点之间的最短路径可以利用Floyd算法代入地图的邻接矩阵求出。这样就成功针对问题本质把地图冒险游戏暂且变成简单无向图上的策略问题。

在地图中，玩家的选择有很多。玩家可能会从起点直奔矿山；也可能直奔村庄；更有甚者可以直接前往终点。但不管玩家的选择多么多样化，根据目的可以大致上分为两种：

其一为前往矿山或者村庄然后开始“挖矿-补给”环节。这种行为的目的在于保证可以安全回到终点的前提下尽量去获取更多来自挖矿带来的收益，但也有可能会出现中途补给不到或者食物短缺的情况，遇到这样的情况时就应该提前选择补给或者撤离。所以很有必要针对这个方向建立模型。

其二为直接奔往终点。当终点离起点较近，矿山却离起点很远的时候，直接前往终点反而是更好的选择。为了考虑周全，这样的情况也有必要讨论。

综上，在抽象化地图后，接下来紧紧围绕两个方向建立模型。

综上所述，保证在规定时间内到达终点的前提下，建立获得尽可能多资金的单目标规划模型如下：

**决策变量**：





**目标函数**：



其中为游戏结束时的资金总量。

**约束条件：**

（1）水和食物重量不能超过负重上限，即



其中为水的箱数，为食物的箱数。

（2）水和食物的初始消耗资金不能超过初始资金，即



（3）玩家依据天气进行行走方案的决策，即



（4）玩家在第天水的剩余量，即



其中为第天水的购买补充量。

（5）玩家在第天食物的剩余量，即



其中为第天食物的购买补充量。

（6）玩家在游戏过程中的收益，即



（7）玩家在村庄补充资源消耗的资金，即



（8）资金总量，即



综上所述，获得尽可能多资金的模型为：





**模型求解：**

### 7.3 模型求解——遗传算法模型

为了求解模型，针对问题的特性我们决定采用了遗传算法。遗传算法简称GA，是以自然选择和生物遗传理论为基础，将生物进化过程中的“物竞天择，适者生存”的规律与群体内部的染色体的随机信息交换机制相结合，是一种高效的全局寻优搜索算法。

尤其是在本题中可能会有走不到终点的情况出现，此时具有“优胜劣汰”特性的遗传算法就非常契合本题的解题思路。并且在使用编码解决问题这点上和30天进行决策有相似之处，因此可以利用二进制编码去代替决策向量。

### 7.3.1 模型建立

遗传算法包括三个基本操作：选择、交叉和变异。如果按照步骤划分，基本过程如下：

**Step1:**计算开始时，以一定数目建立个体，随机地初始化种群；

**Step2:**计算每个个体的适应度函数，产生第一代；

**Step3:**如果不满足优化准则，按适应度选择个体，父代要求基因重组（交叉）

而产生子代；

**Step4:**所有的子代都按照一定概率变异

**Step5:**重新计算子代的适应度函数；

**Step6:**子代被插入到种群中将部分父代取而代之，构成新一代，这一过程一直到设定的满足优化标准为止。然而本次遗传算法的个体编码表示的是决策向量，在计算适应度函数上要加上部分模拟过程，具体流程图如下图所示：

图6.3.1：遗传算法流程图

### 7.3.2 针对问题和模型的改进遗传算法的细节

1）产生初始种群

种群的每个个体的编码代表的是30天内玩家可能会使用的决策向量，根据目标模型可知虽然整个过程分为三段，但决策变量一直是0-1变量。因此在产生种群上我们直接采用长度为30的二进制编码即可。并且因为以赚钱和不在路上浪费过多时间为目标，我们将1的生成概率设为0.8,而0的生成概率为0.2。

2）适应度函数

根据模型和种群内容，最优目标应该是成功完成游戏并且剩余金钱最高的个体。因此我们修改为以输出的剩余金钱变量为主要的衡量对象。并且保存并输出

最好情况的个体编码。

3）选择

针对本题的特点来设置，选择将使适应度较大(即结束游戏后金钱最多)个体有较大的存在机会，反之个体继续存在的机会较小。选择上采用赌轮选择机制，令表示群体的适应度值之总和，其中表示种群中第个染色体的适应度值，这样它产生后代的能力正好为其适应度值所占份额。

4）交叉

根据个体编码的“1多0少”的特点，采用部分匹配交叉()方法，随机选取两个交叉点，根据两个父个体中两个交叉点之间的中间段给出的映射关系生成两个子个体。这样的好处是避免有规律的交叉可能会造成少数地位的“0”规律出现某几个位置，进而导致容易收敛到局部解。

5）变异

由于基于二进值编码的变异操作不能够由简单的变量的翻转来实现，可以利用“1多0少”，随机的在这个编码序列选取两个位置，然后交换他们的位置。这样就实现了个体编码的变异。

六、问题二模型建立与求解

6.1 模型准备

在第三关和第四关中不考虑天气极端的情况下，玩家顺利从起点到终点的天数远小于整个游戏时间段的天数。因此矿山是个必经之地，由此可以将整个游戏流程大体分为两个时间段：第一阶段是从起点到矿山的时间段，第二阶段是到了矿山之后的时间段。不仅如此，两个阶段玩家考虑的问题也不一样。玩家在第一阶段只要考虑是否行走等问题，而在第二阶段玩家要考虑是否要挖矿，是否要补给以及是否要去终点等问题。因此，将整个游戏流程分为两个时间段具有较好的合理性。

6.2 模型建立

#### 6.2.1 第一阶段

第一阶段是从起点到矿山的时间段，在这个时间段玩家要考虑时间风险因素，物质风险因素以及收益风险因素。由此我们建立时间风险指标、物质风险指标以及收益风险指标三个风险指标，助于玩家更好地决策自己的行动方案。

**时间风险指标：**

其中代表玩家剩余的天数，代表在当前位置玩家到达终点的最快天数。时间风险指标越接近数值1，说明玩家剩余时间的可操作性变小，代表玩家游戏失败的风险性越高。

**物质风险指标：**

其中代表玩家剩余的生活必需品，代表玩家期望的生活必需品量。物质风险指标越接近数值1，说明玩家剩余的生活必需品越不足以支撑到达终点，代表玩家游戏失败的风险性越高。

**收益风险指标：**

其中代表玩家的期望收益，代表玩家的等待收益。

综合三个风险指标，最终得到生存风险函数：。由此玩家可以借助生存风险函数得到生存风险值，从而根据玩家的心理风险阈值对自己的行动方案进行决策。其中，决策函数为：



#### 6.2.2 第二阶段

第二阶段是玩家到了矿山之后的时间段，玩家在这个时间段内要考虑是否要挖矿，是否要补给以及是否要去终点等问题，由此我们需要建立三个决策函数对相应状态进行决策，从而求解得一般情况下玩家的最佳策略。

**挖矿决策函数：**

其中当玩家一天的期待收益与挖矿的消耗资源比值小于1则玩家应该放弃挖矿；当比值大于1，玩家应该继续挖矿，从而扩大资源优势。

**返回决策函数：**

其中代表玩家从矿山到村庄再返回矿山所需要的天数。当折返所需要的天数与玩家剩余的天数比值大于1则玩家应该考虑返回终点；当比值小于1，说明玩家的时间充盈，则玩家需要考虑是否可以去挖矿或者去村庄补给。

**补给决策函数：**



其中代表玩家从矿山到村庄所需要的补给，代表玩家在村庄补给所需要的资金，代表玩家在矿山与村庄折返所需要补给消耗的资金。当玩家剩余补给与需要补给的比值大于1则不需要补给，当小于1需要补给且玩家需要考虑补给得值不值问题。当玩家在村庄购买补给消耗的资金与在折返路上消耗的资金之和与玩家在剩余天数所期望收益的比值大于1，则玩家不应该折返矿山而应返回终点。

综合第一和第二阶段，最终可以得到资金总量的目标函数：



其中代表玩家补给的次数，代表玩家挖矿的天数。

七、问题三模型建立与求解

7.1 模型准备

在第五关中事先知道在整个游戏时段内每天天气状况，因此可以在问题一建立的模型基础之上加以改进。现有多名玩家，根据游戏的不同设定，为了保证每名玩家的平均资金总量最大化，每名玩家要减少与其他玩家的接触时间。由此，每名玩家要采用博弈论来决策自己的行动方案，从而使个人的资金总量合理化。

7.2 模型建立

记其中一名玩家的策略为，即当这名玩家认为其他玩家造成的干扰度为，则这名玩家采取策略。对于任意给定的，玩家的策略应该使个人的资金总量最大化，所以满足：



其中，，表示条件下的条件期望，表示事件的概率。

假设每名玩家的相互干扰度分别是资金总量对每名玩家的线性函数，最终可以表示为：



在Nash均衡状态下，所有策略组合满足公式（max），可以使多名玩家的资金总量达到一个均衡的状态。因此可以将问题一的模型改进为：



7.3模型求解——改进遗传算法

7.3.1求解策略分析

该模型的优化求解背景仍然与问题一相似，知道未来的天气，并且玩家人数增加到n人，游戏日期也缩短到10天。虽然当不同的玩家做相同的某些事的时候收益和损益会彼此相互影响，但知道未来天气即确定了天气发生的概率，该问题的本质上仍然是n名玩家的决策向量组合之间所产生的综合效益的优化问题。因此可以考虑改进问题一中的遗传算法，使其可以适应新模型的求解。

7.3.2针对问题的模型改进

具体改进主要体现的三个方面：

第一个方面是在种群建立之前，通过对地图的预处理，选取从起点到矿山的多条路线建立初始路线集合。因为问题中当玩家从一个区域一起向另一个区域前进的时候消耗会变大，因此有必要整理初始路线以方便初始路线的建立和筛选。在具体实现的时候，模拟的玩家先根据自身在初始路线集合中选择一条，这样可以大大减少寻优空间大小。

第二个方面是改进编码，使得可以在一段编码中可以同时体现未来的决策指令向量和初始路线的选择。所以我采用了格雷码和二进制编码相结合的方式，同时也发挥了格雷码的优势，即：交叉变异的截断点如果在格雷码范围内，之后个体的结构和选择也不会因此发生剧烈变化。编码长度调整为13位，格雷码为前三位。

第三个方面为适应度值的计算。将n位玩家在一个地图到达终点后的平均资金作为新的标准。因为不是每一个玩家场场都能成为“领头羊”，所以将平均资金作为衡量整体的标准很合理。

7.3.3流程图

7,4第五关求解结果及模型检验

7.4.1求解结果

针对第五关，我们通过两种情景，以迭代次数30，种群数量40求解出两种情境下的解如下：

1）以前往矿山赚钱为目的

|  |  |  |
| --- | --- | --- |
| 玩家 | 路线 | 余额 |
| 玩家一 | 1->4->4->3->9->9->9(挖)->11->13 | 8975 |
| 玩家二 | 1->2->3->9->9(挖)->9(挖)->11-> | 9200 |

2）以直接前往终点为目的

|  |  |  |
| --- | --- | --- |
| 玩家 | 路线 | 余额 |
| 玩家一 | 1->5->5->6->13 | 9565 |
| 玩家二 | 1->4->4->7->12-13 | 9425 |

显然此时以直接前往终点为目的去求解第五关，比前往矿山更好。

模型准备：

由于多人游戏规则的限制，在未来天气未知的情况下需要尽量避免与其他玩家接触，因此不能单一基于模型二进行求解，还需考虑



八、模型评价与推广

对于问题一，通过对问题的相关条件进行分析，逐步转化为模型的约束条件，使得模型在求解过程中始终保持正确的区间，保证了结果的正确性。同时为了方便求解，采用改进的遗传算法对模型一进行求解，不仅使得求解速度加快，而且最大程度上接近全局最优解，使得解出的游戏方案效果较好。未来可以考虑使用蚁群算法等其他智能算法进行求解。

对于问题二，在未来天气未知的情况下，本模型将游戏分为两阶段并在不同阶段有不同的解决方案同时在每个阶段都考虑到了各种决策条件。从而保证了模型的全面性与稳健性。

在对第三关和第四关求解过程中，采用计算机仿真的方法随机生成不同状况的天气，在经过大量的模拟后统计出了在不同心理阈值情况下的成功率，可以给玩家提供一个完善的游戏策略。

对于问题三，在多人游戏中，在模型一的基础上本模型利用博弈论原理提前规划好不同玩家的游戏路线从而有效的减少了玩家接触情况，并且使得所有玩家收益达到均衡状态，因此该模型在多人游戏领域有着较好的适用性与应用空间。

对于问题四，为了能够在未知天气的情况下以最好的方式完成游戏，本模型基于模型二，在现有决策方式上增加动态博弈模型，使得玩家在每次做计划时都能尽量避免玩家相遇情况，因此本模型能够很好的适应该类问题，并可以预测玩家行为，在其他游戏中也可以得到应用。

九、参考文献

[1]姜启源，谢金星，叶俊.数学模型(第五版）[M].北京：高等教育出版社，2018.369-398

[1]钱小冬.多阶段不确定最优控制的微分动态规划法[D].江苏:南京理工大学,2014. DOI:10.7666/d.Y2521823.

[2]陈晓艳,张东洋,苏学斌, 等.基于改进遗传算法和多目标决策的货位优化策略[J].天津科技大学学报,2020,35(4):75-80. DOI:10.13364/j.issn.1672-6510.20190109.

十、附录