穿越沙漠

摘要

在穿越沙漠游戏中，玩家需要保证自己活着的前提下，尽可能在出沙漠时持有较多的钱。在沙漠中除了普通的行走操作外，还可以停留、购买物资（在村庄处）和挖矿（在矿场处），而不同沙漠中的地形和天气千变万化，还有可能遇到其他玩家，在此复杂的情况下，我们团队做了以下三点工作指导玩家做出最优决策。

在第一问情况下，我们首先计算出玩家每天所能采取的四种行动——行走、停留、挖矿、村庄购物，给带来的玩家状态（天数、持有水、食物和钱）的改变，运用动态规划的思想，通过寻找玩家由某一特定状态能转化的所有可能状态，完成对所有玩家的状态的子状态拆解，然后对其进行遍历，寻找到达最终状态的子状态中金钱数最大的状态，并依次向前寻找。同时在具体的第一关和第二关中，我们运用图论思想简化了考虑的范围，进而大幅提高了计算速度。最终我们得到了第一关和第二关的最优结果，并将其填写在了表格当中。

在第二问未知天气的情况下，我们运用转移矩阵预测了可能的未来天气，运用第一问的算法求出玩家的最优路径后，我们为玩家制定了安全生存法则，使其能够在最坏的天气序列下仍能安全走出沙漠，并规划了每一点玩家到村庄和到终点的玩家的最速撤离路线，综合考虑利益和风险后，对玩家策略做出指导。并测试了在不同随机天气序列下，第三关和第四关的最优行走路线。

在第三问中，由于玩家间的决策会相互影响，我们将每名玩家的目的视作个人利益最大化，通过考虑玩家的备选策略对其它玩家收益的影响，我们穷举了 n 名玩家的策略组合，并根据此矩阵的纳什均衡类型，制定不同的策略指导。若矩阵仅存在一个纯策略纳什均衡点，则所有玩家均采用该策略。若矩阵存在多个均衡点，我们采用自定义的 Pareto 上策平衡点作为最优决策，若矩阵仅存在混合策略均衡点，则我们认为玩家将采用概率空间中最大概率的决策。并据此解决了第五关和第六关。

综上，我们能有效求解未来天气序列已知时玩家的最佳策略。而在未来天气序列未知时，我们可以保障玩家安全的情况下，让玩家获取尽可能大的收益。同时在玩家数大于 1 时，我们可以基于博弈论给出每名玩家的最理性解。

关键字： 动态规划　安全生存法则　纳什均衡

目录

一、 问题重述 **3**
1.1 问题的提出 3
二、 模型的假设 **3**
三、 符号说明 **3**
四、 模型的建立 **3**
4.1 问题一模型 3
4.1.1 数据及题意分析 4
4.1.2 提取关键路径 4
4.1.3 递推方程 6
4.2 问题二模型 8
4.2.1 玩家资产模型 8
4.2.2 后续天气预测 9
4.2.3 玩家决策模型 10
4.3 问题三模型 12
4.3.1 基于博弈论的最优策略 12
五、 模型的求解及结果分析 **14**
5.1 问题一模型求解 14
5.2 问题二模型求解 15
5.2.1 简化关卡地图 15
5.2.2 第三关求解 15
5.2.3 第四关求解 16
5.3 问题三模型求解 18
5.3.1 第五关求解 18
5.3.2 第六关求解 18
六、 模型的优缺点 **19**
6.1 模型的优点 19
6.2 模型的缺点 19
七、 附件 **20**

一、 问题重述

穿越沙漠小游戏要保证玩家在规定时间内走出沙漠，并使游戏结束后的钱尽可能多。玩家有水和食物两种物资，可在起点处购买，不同物资有不同的重量和价格，且玩家负重上限为 1200kg。在不同天气和决策下会消耗不同数量的物资，在村庄中玩家可以高价购买物资，若物资耗尽且无法及时得到补充则视作游戏失败。在矿山处玩家可以消耗一天时间挖矿以获得金钱，但消耗量是基础消耗的三倍。如果多名玩家同时作出某决策，如下矿或购物或沿同一道路前进，则决策成本会提高。

1.1 问题的提出

问题一：在只有一名玩家且每天天气已知的情况下，试分别给出附件中第一、二关的最佳路径。

问题二：在只有一名玩家且每天天气仅在当天才可知的情况下，试分别给出附件中第三、四关的最佳路径。

问题三：

（1）存在有多名玩家且每天天气已知，每名玩家的行进路线在第 0 天做出后即无法更改，试给出各玩家第五关的行进路线。

（2）存在有多名玩家且每天天气仅在当天才可知，每名玩家结束后可以知道其他玩家剩余资源量，试给出第六关各玩家应采取的策略。

二、 模型的假设

- 购买食物仅能按整数箱购买。
- 当存在多条最短路径时，我们将所有最短路径视作等价的并仅考虑一条。
- 问题中提供的天气序列均为该地天气。
- 玩家在想办法获得更多利益时，首先应当考虑自身安全。
- 当有多名玩家时，每名玩家均会采取对自己最有利的策略。

三、 符号说明

四、 模型的建立

4.1 问题一模型

在已知全部天气情况下规划行进路线使之最终剩余的钱最多，即求解所有路线中的最优解，因此本题可以看作一个全局优化问题。在此基础上，我们对关卡作以下抽象化

表 1　本文所用符号的相关说明与解释

符号	意义
Ω_t	第 t 天天气
C_f	基本食物消耗量
C_w	基本水消耗量
M_t	第 t 天玩家状态向量
X_t	第 t 天玩家位置
$,m_t$	第 t 天玩家持有金钱
w_t	第 t 天玩家
f_t	第 j 个任务 (包) 的定价
λ_{12}	第 j 个任务对第 i 个会员完成概率的距离影响因子
ϵ_i	第 i 个会员的信誉值
σj	第 j 个任务完成概率的定价影响因子

处理。

4.1.1 数据及题意分析

根据第六条游戏规则，玩家到达终点后可退回剩余的水和食物，每箱退回价格为基准价格的一半。考虑到无论在起点还是在村庄购买食物，购买的量都是可以自定义的，且初始状态下（不购买时）玩家持有的食物和水量都是零，因此有**定理 1**:

定理 1 对于任意一策略，如果其在终点存在多余的食物或水，可以通过在最后一次购买食物或水时，选择少购买相应量的食物或水，达到更优的策略。

推论 1 最佳策略中，终点处剩余的水和食物均为零。

4.1.2 提取关键路径

首先，若做出行走决策，玩家每次只能跨越相邻的地块，故不妨将每个地块看作图论中的顶点，而相邻地块代表的节点间存在距离为 1（天）的边，我们即可将关卡抽象成一个赋权图 $G=(V,E,A_0)$，其中 V 是顶点集，在顶点集 V 中，有四类特殊顶点 V_s:

起点、村庄、矿场、终点。E 是边集，其中：

$$A_0 = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

代表邻接矩阵，其中：

$$a_{ij} = \begin{cases} 1, \text{如果 i、j 相邻} \\ \infty, \text{如果 i、j 不相邻} \end{cases}$$

由于本题从节点 i 到节点 j 的时间成本是一样的，因此可以视作无向图，即：$a_{ij} = a_{ji}$

根据题目要求，玩家存在最迟到达终点时间，且需要尽可能让结束时自己有更多的钱，而食物和水的物资是由钱购买的，因此我们有**定理 2**：

定理 2 最佳决策中玩家不会做既会拖延时间又不能带来利益的决策。

在问题一中，最佳决策方案一开始就能确定，在此条件下行走于特殊节点间时，若作出了不走特殊节点间最短路径的决策，不仅会浪费一天时间，还会白白消耗玩家自身的食物和水资源，无论在节约时间层面还是在经济层面均是不利条件。因此我们可以得到**推论 2**。

推论 2 在特殊节点间步行时，玩家不会走非最短路径的路线。

因此，在每个关卡中，并非所有路径都是有效的，通过 R.W.Floyd 算法计算各个特殊节点间的最短路径，我们可以将各个关卡简化为仅含最短路径的简化关卡。算法基于邻接矩阵 A_0，具体流程如下：

- Floyd 算法利用递推产生一个矩阵序列 $A_1, A_2 \dots A_n$，其中矩阵 A_k 第 i 行第 j 列元素 $A_k(i,j)$ 表示从顶点 v_i 到顶点 v_j 路径上所经过的顶点序号不大于 k 的最短路径长度。
- 对于每个矩阵 A_k，我们用矩阵 A_{k-1} 对其进行迭代优化，迭代公式为：

$$A_k(i,j) = min(A_{k-1}(i,k) + A_{k-1}(k,j)) \tag{1}$$

  其中，k 是迭代次数，$i, j, k = 1, 2, \dots n$
- 当 $k = n$ 时，迭代完成，输出的 A_n 即所求的最短路径矩阵，A_{ij} 代表从 i 到 j 的最短路径权值。我们分别对第一关和第二关的特殊节点计算最短路径权值，并输出其最短路径行走方式，将其标注在原图上（最短路径不止一条时，仅标注其中之一），如下**图 1**和**图 2**所示。

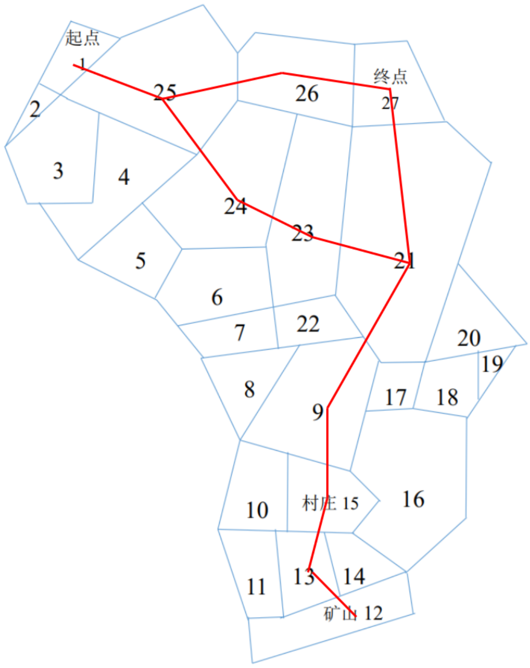

图 1　第一关简化图

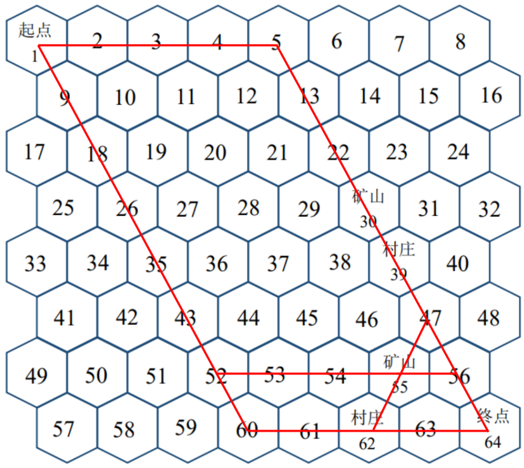

图 2　第二关简化图

此处简化路径的含义是，即便删减图中沙漠区块，仅留下图中红线经过的区块，也不会影响后续对玩家最佳行进策略的分析。

4.1.3 递推方程

设第 t 天的天气情况为 $\Omega_t \in \{\Omega_{sun}, \Omega_{heat}, \Omega_{sand}\}$，其中 Ω_{sun} 代表晴朗天气，Ω_{heat} 代表高温天气，Ω_{sand} 代表沙暴天气。同时，设在三种天气下的食物的消耗基本量分别为:

$$C_f = \begin{cases} C_f^{\Omega_{sun}}, \text{晴朗天气时} \\ C_f^{\Omega_{heat}}, \text{高温天气时} \\ C_f^{\Omega_{sand}}, \text{沙暴天气时} \end{cases}$$

水的消耗基本量分别为:

$$C_w = \begin{cases} C_w^{\Omega_{sun}}, \text{晴朗天气时} \\ C_w^{\Omega_{heat}}, \text{高温天气时} \\ C_w^{\Omega_{sand}}, \text{沙暴天气时} \end{cases}$$

则第 t 天的食物和水的消耗基本量可以用 $C_f^{\Omega_t}, C_w^{\Omega_t}$ 表示。为了表示玩家在各天的状态，包括所在地、储备水等信息，我们引入玩家在第 t 天结算时的状态向量 $M_t = (t, X_t, w_t.f_t)$，

其中四个量分别代表玩家现在的时间第 t 天、该天结束时的所在地、该天结束时的储备水和该天结束时的储备食物。

接下来我们将分五种情况讨论第 t 天玩家做不同决策时，每日结算时玩家的状态向量。

- 玩家停留。玩家状态变化可用下式表示：

$$\begin{cases} x_{t+1} = x_t \\ w_{t+1} = w_t - C_w^{\Omega_t} \\ f_{t+1} = f_t - C_f^{\Omega_t} \\ m_{t+1} = m_t \end{cases} \tag{2}$$

  我们用一个函数定义此类状态变化：$Case_1(M_t, m_t) = (M_{t+1}, m_{t+1})$

- 玩家行走。此情况发生的必要条件是玩家不遇到沙暴，即 $\Omega_t \neq \Omega_{sand}$。由于行走情况下消耗的物资为基本物资的两倍，因此玩家状态变化可用下式表示：

$$\begin{cases} x_{t+1} \in \{x_t\text{所在邻域}\} / x_t \\ w_{t+1} = w_t - 2C_w^{\Omega_t} \\ f_{t+1} = f_t - 2C_f^{\Omega_t} \\ m_{t+1} = m_t \end{cases} \tag{3}$$

  我们用一个函数定义此类状态变化：$Case_2(M_t, m_t) = (M_{t+1}, m_{t+1})$

- 玩家挖矿。此情况发生的必要条件是玩家第 t 天所在地点为矿场，即 $x_{t+1} = x_t = x_{\text{矿场}}$。

$$\begin{cases} x_{t+1} = x_t = x_{\text{矿场}} \\ w_{t+1} = w_t - 3C_w^{\Omega_t} \\ f_{t+1} = f_t - 3C_f^{\Omega_t} \\ m_{t+1} = m_t \end{cases} \tag{4}$$

  我们用一个函数定义此类状态变化：$Case_3(M_t, m_t) = (M_{t+1}, m_{t+1})$

- 玩家途径村庄购买物资并停留。此情况发生的必要条件是玩家第 t 天所在地点为村庄，即 $x_{t+1} = x_t = x_{\text{村庄}}$。

$$
\begin{cases}
x_{t+1} = x_t = x_{\text{村庄}} \\
w_{t+1} = w_t - C_w^{\Omega_t} + B_{wt} \\
f_{t+1} = f_t - C_f^{\Omega_t} + B_{ft} \\
m_{t+1} = m_t
\end{cases} \tag{5}
$$

我们用一个函数定义此类状态变化：$Case_4(M_t, m_t) = (M_{t+1}, m_{t+1})$

- 玩家途径村庄购买物资不停留。此情况发生的必要条件是玩家第 t+1 天所在地点为村庄，即 $x_{t+1} = x_{\text{村庄}}$。

$$
\begin{cases}
x_{t+1} = x_t = x_{\text{村庄}} \\
w_{t+1} = w_t - 2C_w^{\Omega_t} + B_{wt} \\
f_{t+1} = f_t - 2C_f^{\Omega_t} + B_{ft} \\
m_{t+1} = m_t
\end{cases} \tag{6}
$$

我们用一个函数定义此类状态变化：$Case_5(M_t, m_t) = (M_{t+1}, m_{t+1})$

且对于上述任意一个状态 M_t，有如下约束：

$$
\begin{cases}
m_t \geq 0 \\
f_t \geq 0 \\
w_t \geq 0 \\
3 * w_t + 2 * f_t \leq 1200
\end{cases} \tag{7}
$$

4.2 问题二模型

在此模型中，由于未来天气是不可知的，玩家的最佳决策会实时发生变化，因此适用于问题一模型的**推论 2**和**推论 1**均不适用，但我们还是对问题做了些许简化，方便我们后续的分析。

4.2.1 玩家资产模型

在第一问的模型中，我们将水、食物等基础物资和玩家持有的钱分开讨论，是因为**推论 1**，最终状态玩家的玩家所持食物和水的量均为零。而在本问的大背景下，由于天气状况不定，玩家无法通过调整购买策略达到终点零物资剩余的目标，而是要在终点进行物资的转化。换言之，我们无法像第一问那样仅根据玩家终点的物资数不为零就判断玩家采取的不是最佳策略。

因此，我们将终点视作特殊的村庄，在最后走到终点的日期，考虑水的消耗的同时进行食物和水的换算，则到终点处的状态向量变化如下：

$$\begin{cases} x_{t+1} = x_t = x_{\text{终点}} \\ w_{t+1} = w_t - 2C_w^{\Omega_t} - S_{wt} = 0 \\ f_{t+1} = f_t - 2C_f^{\Omega_t} - S_{ft} = 0 \\ m_{t+1} = m_t + 2.5 * S_{wt} + 5 * S_{ft} \end{cases} \tag{8}$$

式中，2.5 和 5 分别为终点处水和食物换算单价。

采取此转换后，**推论 1**将继续成立，我们在第一问中采用的动态规划算法也可以顺利使用。

4.2.2 后续天气预测

通过研究第一关、第二关和第五关的天气我们发现，在 70 天的天气中，共出现 23 次晴朗天气、35 次高温天气和 12 次沙暴天气，晴朗天气变成晴朗天气共出现 7 次，晴朗天气变成高温天气共出现 12 次，晴朗天气变成沙暴天气共出现 4 次，高温天气变成晴朗天气共出现 11 次，高温天气变成高温天气共出现 15 次，高温天气变成沙暴天气共出现 6 次，沙暴天气变成晴朗天气共出现 4 次，沙暴天气变成高温天气共出现 6 次，沙暴天气变成沙暴天气共出现 2 次。

我们假设第 n 天天气仅与第 n-1 天有关而与过去的天气无关，据此我们可以用转移矩阵大致表示天气变化的规率，由上述数据我们可得不同天气间的转移矩阵为 A_c：

$$A_c = \begin{bmatrix} \frac{7}{23} & \frac{11}{32} & \frac{4}{12} \\ \frac{12}{23} & \frac{15}{32} & \frac{6}{12} \\ \frac{4}{23} & \frac{6}{32} & \frac{2}{12} \end{bmatrix}$$

式中，$i, j = 1, 2, 3$ 分别代表晴朗天气天、高温天气和沙暴天气，A_{ij} 表示 i 对应的天气后转化到 j 对应的天气的可能性。

同时我们也可以得到不同天气的出现概率：

$$
\begin{cases}
P_{sun} = \dfrac{23}{67} \\[2em]
P_{heat} = \dfrac{35}{67} \\[2em]
P_{sand} = \dfrac{12}{67}
\end{cases}
\tag{9}
$$

如果我们知道了第 t 天的天气 w_t，则根据转移矩阵的定义，我们可以用下式表达第 t+n 天的各天气出现的概率：

$$
\begin{bmatrix} p_{sun}^{t+n} \\ p_{heat}^{t+n} \\ p_{sand}^{t+n} \end{bmatrix} = A_c^n \cdot \begin{bmatrix} p_{sun}^{t} \\ p_{heat}^{t} \\ p_{sand}^{t} \end{bmatrix}
\tag{10}
$$

其中，$\begin{bmatrix} p_{sun}^{t} \\ p_{heat}^{t} \\ p_{sand}^{t} \end{bmatrix}$ 是一个 3*1 的向量，第 1、2、3 行分别代表晴朗天气、高温天气、沙暴天气出现的概率，当已知第 t 天天气为晴朗时，则有：

$$
\overrightarrow{w_t} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}
\tag{11}
$$

用 1 代表晴朗天气已出现。

4.2.3 玩家决策模型

我们的首要任务是保证玩家的安全，不会因为缺少物资死亡。因此我们指导玩家做出的决策将首先考虑决策的安全性，再考虑决策带来的利益。明确了此标准后，一个用当天天气指导决策的模型就跃然纸上了。

对于任意时刻 t、任意位置的玩家，我们首先利用转移矩阵，根据当天天气得到后续天气的概率预测，然后根据下式计算出第 t+i 天基础物资消耗量的期望 $\overline{C_f^{t+i}}$、$\overline{C_w^{t+i}}$：

$$
\begin{bmatrix} \overline{C_f} \\ \overline{C_w} \end{bmatrix} = \begin{bmatrix} C_f^{\Omega_{sun}} & C_f^{\Omega_{hwat}} & C_f^{\Omega_{sand}} \\ C_w^{\Omega_{sun}} & C_w^{\Omega_{hwat}} & C_w^{\Omega_{sand}} \end{bmatrix} \cdot \begin{bmatrix} p_{sun}^{t+i} \\ p_{heat}^{t+i} \\ p_{sand}^{t+i} \end{bmatrix}
\tag{12}
$$

根据求出的各天期望物资消耗量，我们同样可采用 4.1.3 的递推方程求出玩家的状态变化，区别仅仅在于将 C_f、C_w 替换为了 $\overline{C_f^{t+i}}$、$\overline{C_w^{t+i}}$。

而为了保障玩家的生存，我们通过下面的方法计算出了每一天的玩家最速撤离条件：

- 玩家处于任意点处，将目的地设为终点，计算最短路径并将其设置为最速撤离路径。此时能保障玩家撤离需要的条件是最坏情况走完此最速撤离路径的最短时间 t_{min}，和最速撤离路径下的最坏情况的物资消耗量 (w_g, f_g)。
  对于第三关，最坏情况即天气全部是高温，而对于第四关，最坏情况即天气全部是高温再加上路上遇到一天沙暴（少量沙暴）。
- 玩家处于任意点处，将目的地设为村庄，计算最短路径并将其设置为最速撤离路径。此时能保障玩家撤离需要的条件是最速撤离路径下的最坏情况的物资消耗量 (w_v, f_v)，和走完此最速撤离路径、到达村庄的最短时间 t_{min1}，再加上村庄到达终点的最短时间 t_{min2}。
- 对于两个最速撤离条件 (w_g, f_g, t_{min}) 和 $(w_v, f_v, t_{min1} + t_{min2})$，只要玩家剩余的物资和时间 w_l, f_l, t_l 仍大于上述任意一个最速撤离条件，我们则认为玩家处于安全状态。由于仅考虑最坏情况，上述两个撤离条件中三个量是与所需天数严格正相关的，因此不会出现最速撤离条件 A 的水需求大于 B，而最速撤离条件 B 的天数需求却大于 A 的情况。

据此，我们采取以下两种玩家安全保障措施：

- 决策策略：对于第 t 天早上知道天气并得出临时最优决策后，我们将考虑按照最优路径决策一天后，身上剩余资源 (w, f, t) 是否满足最速撤离条件。若是，则按照最优路径行走，并规划第 t+1 天的最速撤离路径；若否，则按照第 t-1 天预测的最速撤离路径离开沙漠。对于 t=0 的时刻的撤离路径，也就是起点处，我们设置为起点到终点的最短路径。
- 购买策略：对于在物资购买点（起点或村庄）的玩家，假设其此时计算的最优策略建议的购买策略是 B_{f0}, B_{w0}（箱），而在此购买点，为了在执行完一天最优策略后还要达到最速撤离条件需要的最小物资需求为 B_{fl}, B_{wl}，我们的购买策略是：

$$B_f = \begin{cases} \max(B_{f0}, B_{fl}), if \ \ 3 * \max(B_{f0}, B_{fl}) + 2 * \max(B_{w0}, B_{wl}) \le 1200 \\ \max(B_{fl}, \lceil \dfrac{1200 - 2 * \max(B_{w0}, B_{wl})}{3} \rceil), if \ \ 3 * \max(B_{f0}, B_{fl}) + 2 * \max(B_{w0}, B_{wl}) > 1200 \end{cases} \tag{13}$$

$$B_w = \begin{cases} \max(B_{w0}, B_{wl}), if \ \ 3 * \max(B_{f0}, B_{fl}) + 2 * \max(B_{w0}, B_{wl}) \le 1200 \\ \max(B_{wl}, \lceil \dfrac{1200 - 3 * \max(B_{f0}, B_{fl})}{2})\rceil, if \ \ 3 * \max(B_{f0}, B_{fl}) + 2 * \max(B_{w0}, B_{wl}) > 1200 \end{cases} \tag{14}$$

上式看着比较复杂，但意义简单明确，我们的策略即是在最优购买策略和最小物资

需求中，对水和食物都选择较大的一方。

值得注意的是，由于原来两个购买策略均不会超过负重上限，若选择完后超出负重上限 1200kg，则必然有 $(B_{f0} - B_{fl})(B_{w0} - B_{wl}) < 0$，此时我们优先保证撤离条件的物资需求，再在没达到最优要求的物资方面尽量去靠近最优购买策略。

4.3 问题三模型

本问中，由于存在多个人决策，且每个人的决策会对其他人产生影响，为了帮助每名玩家做出对自己的利益最大的决策，我们采用博弈论对本问题建模。

4.3.1 基于博弈论的最优策略

在天气已知的情况下，对于参与游戏的 n 名玩家，其在起点处便可做出整场游戏的策略，因此对于此类问题，我们可以视作 n 名玩家进行了一次静态博弈。

对于玩家的潜在可能策略，我们采用以下流程获得：

- 根据 4.1 的方法，计算出在当前天气序列下最优的单人策略 T_1。
- 考虑当一个人已经以 T_1 路线行走后，重新调整各个时间段下各个线路的花费，计算出此时的最短路径 T_2
- 按照上述方法，加入人数到第 k 个人，其中 k 为后期设置的参数，表示我们的预备策略数。
- 重新计算第一个人的最佳策略，更新 T_1，同理，更新 T_2、T_3⋯⋯
- 迭代若干次上述过程，得到我们的预备策略集合：$\{T_1, T_2, \ldots T_k\}$。

我们假设在起点处，玩家共有 k 种策略 $\{T_1, T_2, \ldots T_k\}$ 选择，当玩家选择的策略集合为 $T = \{T_{i1}, T_{i2}, \ldots T_{in}\}$ 时($\{i1, i2, \ldots in\} \in \{1, 2, \ldots k\}$),各玩家的收益为 $\left\{m_T^1, m_T^2, \ldots m_T^n\right\}$。然后我们分下述三种情况求解 n 名玩家们做出的最优策略集合 T：

- 策略集合 T 存在纯策略且单一的纳什均衡。即存在一个策略 T_i，每名玩家在无论其他玩家做什么决策的情况下，选择策略 T_i 的利益都大于其他策略的利益。此时根据博弈论，玩家应当自发倾向选择此策略，此时玩家策略集合 $T = \{T_i, \ldots T_i\}$。
- 策略集合 T 存在纯策略多个纳什均衡，每个人的策略取决于其他参与人的策略，以概率分布 P 随机地选择不同的策略，此时我们采用帕累托上策均衡作为我们的选择，不过不同的是，除了考虑均值最大，我们还考虑了决策对各玩家的公平性。我们采用下式作为各个策略的评分：

$$f(T) = \alpha * \bar{m_T} + (1 - \alpha) * \delta(m_T) \tag{15}$$

式中，$\bar{m_T}$ 代表 n 名玩家做出的策略集合为 T 时收益的均值，δm_T 代表 n 名玩家做出的策略集合为 T 时收益的标准差。α 代表玩家自定义的权重，当玩家更看重集体

利益最大化时，应当选择较大的 α；反之，当玩家更看重个人利益分配均衡时，应当选择较小的 α。算出各个策略组合的得分后，我们选出所有组合中得分最高的组合 T，作为 n 名玩家的策略。

- 策略集合 T 仅存在混合策略的纳什均衡。

  我们假设各玩家均以某种概率分布随机地选择不同的策略，对于任意一名玩家 J 而言，这样的概率选择应当使他不管其余 n-1 名玩家做出什么样的策略，他的期望收益都是相等的。设满足上条件的概率选择为 $P=\{P_1,P_2,...P_n\}$。则应有：

$$\begin{cases} m^J_{TT_1}*P_1+m^J_{TT_2}*P_2+...+m^J_{TT_n}*P_n\equiv C \\ P_1+P_2+...+P_n=1 \end{cases} \tag{16}$$

  其中 $\{TT_1,TT_2,...TT_k\}$ 代表玩家 J 分别做出 $\{T_1,T_2,...T_k\}$ 选择时，n 名玩家集体的策略集合；C 是某常数。我们假设对于 n 名玩家，其均会采用我们提供的混合策略方法进行决策，则此 n 名玩家是等价的，因此不同玩家的概率选择 $P=\{P_1,P_2,...P_n\}$ 应当也是一样的，且其期望收益 C 也应当是一样的。因此有：

$$\begin{cases} m^1_{TT_1}*P_1+m^1_{TT_2}*P_2+...+m^1_{TT_n}*P_n\equiv C_1 \\ m^2_{TT_1}*P_1+m^2_{TT_2}*P_2+...+m^2_{TT_n}*P_n\equiv C_2 \\ \vdots \\ m^n_{TT_1}*P_1+m^n_{TT_2}*P_2+...+m^n_{TT_n}*P_n\equiv C_n \\ P_1+P_2+...+P_n=1 \\ C_1=C_2=...=C_n \end{cases} \tag{17}$$

  求解上述线性方程组后，我们即可用 k 种策略在不同情况下的成本算得每名玩家的选择概率集合 P：

$$\begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix} = \begin{bmatrix} m^1_{TT_1} & m^1_{TT_2} & \dots & m^1_{TT_n} \\ m^2_{TT_1} & m^2_{TT_2} & \dots & m^2_{TT_n} \\ \vdots & \vdots & \ddots & \vdots \\ m^n_{TT_1} & m^n_{TT_2} & \dots & m^n_{TT_n} \end{bmatrix}^{-1} \cdot \begin{bmatrix} C \\ C \\ \vdots \\ C \end{bmatrix} \tag{18}$$

  此时，我们即得到了纳什均衡下，玩家做出的最优混合策略。不过将概率代入模型进行计算会引入许多不确定性，因此我们近似认为每名玩家会选取拥有最大概率的策略作为自己的最佳策略，其集合为 T。

而在天气未知时，玩家无法在起点处就做出全部策略，因此我们采用 4.2 的模型思想，对于任意时刻 t、任意位置的玩家，我们首先利用转移矩阵，根据当天天气得到后续天气的概率预测，然后计算出第 t+i 天基础物资消耗量的期望，据此计算出每名玩家

在单人模式下应采取的最优策略，注意此策略在多人模式并不一定最优，我们仅仅是将其作为策略考虑的一种。

接着，我们按照与本节之前提到的预备策略产生方式相似的方式：

- 根据 4.1 的方法，计算出在当前天气序列下最优的单人策略 T_1^1。
- 考虑当一个人已经以 T_1^1 路线行走后，重新调整各个时间段下各个线路的花费，计算出此时的最短路径 T_1^2
- 按照上述方法，加入人数到第 k 个人，其中 k 为后期设置的参数，表示我们的预备策略数。
- 重新计算第一个人的最佳策略，得到 T_2^1，同理，更新 T_2^2、$T_2^3 \cdots\cdots$
- 迭代若干次上述过程，得到第 i 名玩家的预备策略集合：$\left\{T_1^i, T_2^i, \ldots T_k^i\right\}$。

可以得到任意时刻 t 下第 i 名玩家的预备策略集 $\left\{T_1^i, T_2^i, \ldots T_k^i\right\}$。同样按照本节之前提到的博弈论思想，我们可以求得玩家在第 t 时刻选择的最优策略 T，本次决策所有玩家即按照 T 进行行动。在每天都进行一次这样的操作，我们即可以求解天气未知的情况下玩家的行动。

五、 模型的求解及结果分析

5.1 问题一模型求解

在上一节中，我们已经成功找到了从玩家第 t 天的状态推导到玩家第 t+1 天状态的表达式，且经过之前的分析，本问题的核心思路是寻找能达到最大金币数的策略，故我们选择动态优化作为我们本问的求解算法。

动态规划算法通常用于求解具有某种最优性质的问题。在这类问题中，可能会有许多可行解。每一个解都对应于一个值，我们希望找到具有最优值的解。其基本思想是将待求解问题分解成若干个子问题，先求解子问题，然后从这些子问题的解得到原问题的解。

本问题中，对于任意一可能的状态 $M_{t+1} = (t+1, j, w_t, f_t)$，我们都可以用 $\left\{Case_1(M_t^1, m_t^1), Case_2(M_t^2, m_t^2), Case_3(M_t^3, m_t^3), Case_4(M_t^4, m_t^4), Case_5(M_t^5, m_t^5)\right\}$

中的一种表示 M_{t+1}，式中：

$$\begin{cases} M_t^1 = (t, j_1, w_t^1, f_t^1) \\ M_t^2 = (t, j_2, w_t^2, f_t^2) \\ M_t^3 = (t, j_3, w_t^3, f_t^3) \\ M_t^4 = (t, j_4, w_t^4, f_t^4) \\ M_t^5 = (t, j_5, w_t^5, f_t^5) \\ \{j_1, j_2, j_3, j_4, j_5\} \in \{j\text{所在邻域}\} \end{cases} \tag{19}$$

而对上述五个 M_{ti} 中的任意一个，我们可以用同样的办法由 M_{t-1} 经过某个 $Case_i$ 变化得到。而在遍历了初始购买的食物和水的量后，我们即可得到一系列玩家第零天的初始状态 $M_0 = \left\{M_0^1, M_0^2, \dots M_0^n\right\}$，其由初始购买不同物资的数量决定，这样我们即可从第 0 天的所有状态递推到第 t 天的所有状态。

若想让状态 M_{t+1} 拥有尽可能大的金钱值，我们应在所有能得到状态 M_{t+1}^n 的状态 M_t^n 中，选取当中具有最大金钱量的状态 M_t^i，作为到状态 M_{t+1}^n 的最佳变化源，即：

$$\begin{aligned} M_{t+1} &= Case_j(M_t^i) \\ s.t. \quad m_t &= \max_{i \in \{1, \dots n\}} Case_j(M_t^i) \end{aligned} \tag{20}$$

其中 $j \in \{1, 2, 3, 4, 5\}$。最后根据**推论 1**可知，对于最佳策略，最终状态应为 $M_t = (t_{end}, j_{end}, 0, 0)$，据此进行动态规划，即可求解出第一关和第二关的结果，如如**图 3**和**图 4**所示，箭头代表行走路线，箭头上的数字代表行走过程发生的日期数。

5.2 问题二模型求解

5.2.1 简化关卡地图

运用与 4.1.2 同样的方法，我们可以将关卡用图论理论表示，但由于玩家可能会由于天气原因，在到达原设定目的地的中途做出改变目的地的决策，我们不再能仅考虑特殊顶点间的最短路径而忽略其他顶点间的最短路径。不过对于不处于任意一特殊顶点间最短路径的顶点，由**推论 2**可知，我们仍可以将其忽略不计。因此我们将第三关和第四关的地图抽象表示，如**图 5**和**图 6**所示。

5.2.2 第三关求解

我们首先根据之前得到的各个天气出现的概率（**公式 9**），随机生成一天气序列，由于第三关限制不出现沙暴天气，我们把随机天气序列中的沙暴全部换成仅次于沙暴的高温天气。

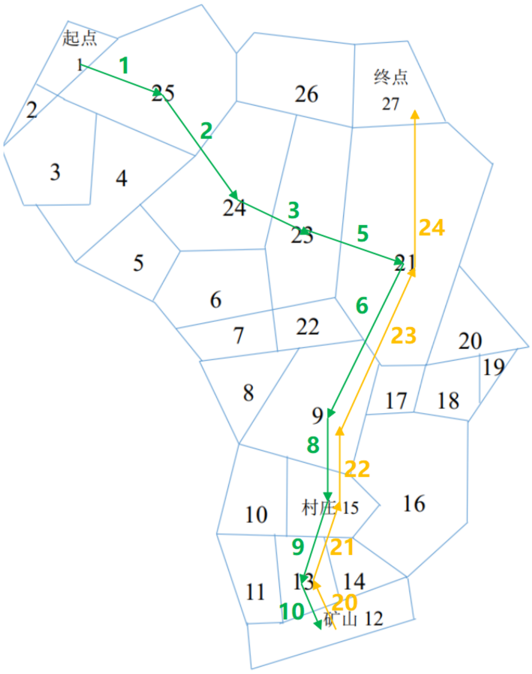

图 3　第一关策略

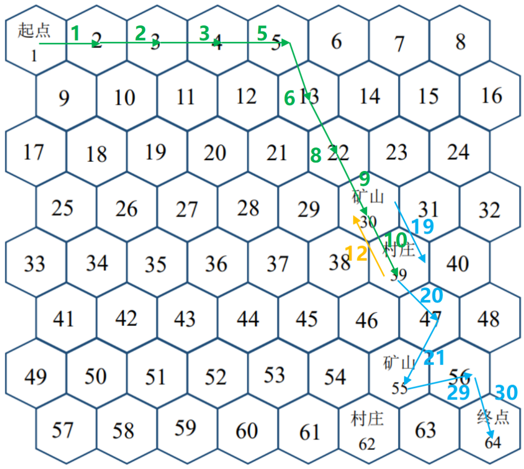

图 4　第二关策略

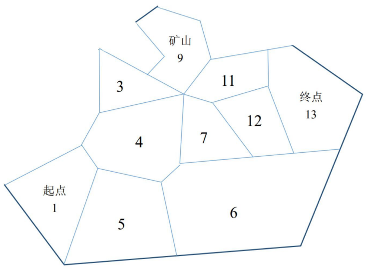

图 5　第三关简化图

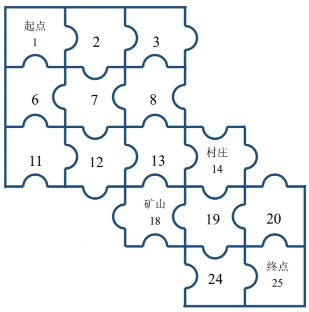

图 6　第四关简化图

通过数千次模拟，我们观察发现，无论出现哪种天气序列，玩家的最优策略都将规划他于第三天或第四天到达终点，我们猜测这可能是因为矿山收益较小、不能弥补成本导致的。而玩家的最终受益也取决于具体天气情况，我们将第三关的路线和几次模拟结果列在**图 7**和**表 3**中。

5.2.3 第四关求解

同理，我们在此关中按照**公式 9**设置随机天气序列，设计的三十天天气和玩家所走路线如下：

(高温, 沙暴, 晴朗, 晴朗, 晴朗, 高温, 高温, 晴朗, 高温, 高温,
高温, 高温, 晴朗, 高温, 高温, 高温, 高温, 高温, 高温, 高温,

表 2　不同天气下第三关收益

天气	行走路线	最终收益
晴-高-晴-高-⋯⋯	1-5-6	9350
高-晴-高-晴-⋯⋯	1-5-6	9270
高-高-高-晴-⋯⋯	1-5-6	9190

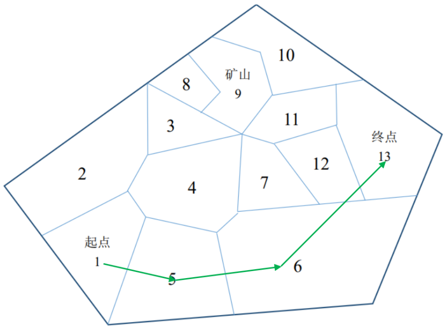

图 7　第三关答案

晴朗, 晴朗, 高温, 晴朗, 高温, 高温, 晴朗, 晴朗, 高温, 高温）

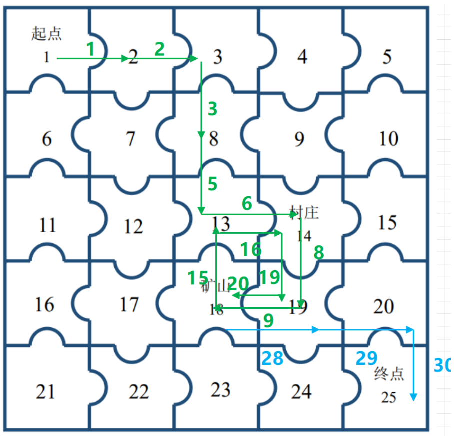

图 8　第四关答案

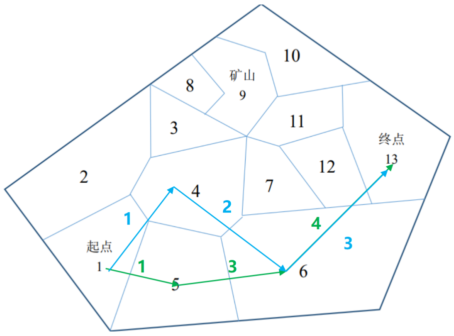

图 9　第五关答案

5.3 问题三模型求解

5.3.1 第五关求解

在该关卡中，我们取玩家策略数 $k=2$，迭代计算得到起点处玩家有如下两种备选策略：我们先取 S1，S4 两种策略进行博弈，在这种情况下，我们可以得到下表：接下来

表 3　博弈结果

玩家 2＼玩家 1	S1	S4
S1	（8185，8185）	（8935，8870）
S4	（8870，8970）	（7740，7740）

计算混合策略下的概率，假设玩家取 S1 策略的概率是 P1，取 S4 策略的概率是 P2，那么玩家 1 要在不论玩家 2 如何选择策略的情况下达到期望收益均等：

$$8185 * P1 + 8935 * P2 = 8870 * P1 + 7740 * P2P1 + P2 = 1 \tag{21}$$

求解得到：

$P1 = 1.0127, P2 = -1.0127$

所以不存在混合策略下的纳什均衡，所以我们采取纯策略下的纳什均衡：即一个人选择策略 S1，另一个人选择策略 S4。

同时我们也计算了其他策略不同组合的博弈结果，发现只有一个人采取策略 S1，另一个人采取策略 S2 时结果才是最优的。

5.3.2 第六关求解

当玩家数等于 3 时，我们准备给每名玩家提供 4 种策略，迭代次数设置为 2，这导致预备最优策略集合的排列组合数量将上升到 4^3 种，而这样的组合数我们将会计算 60

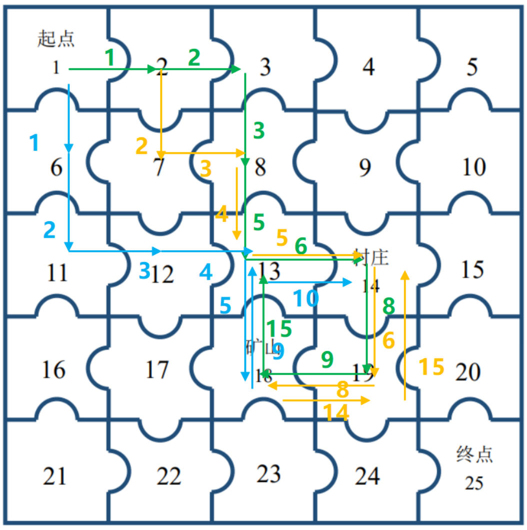

图 10　第六关答案

次，也就是 3840 种，这相当于我们将会运行数千次第一问的优化算法，数量过于庞大，无法进行求解。因此我们降低了提供给玩家的策略，并将迭代次数设置为 1，并采用与第四问相同的设定天气序列，求解出的路径如图 10

为了让结果更清晰，本图仅展示前十五天结果，三种颜色的线条分别代表三名玩家的行动决策。

六、 模型的优缺点

6.1 模型的优点

1. 本文能有效求解未来天气序列已知时玩家的最佳策略。
2. 在未来天气序列未知时，本文可以保障玩家安全的情况下，让玩家获取尽可能大的收益。
3. 在玩家数大于 1 时，本文可以基于博弈论给出每名玩家的最理性解。

6.2 模型的缺点

1. 在未知天气序列时，本文预测所用数据不充足。
2. 本文可能过分看重玩家安全性，导致损失了一部分收益。
3. 第三问计算量太大。

七、 附件