赛区评阅编号（由赛区组委会填写）：

2020 高教社杯全国大学生数学建模竞赛

承　诺　书

我们仔细阅读了《全国大学生数学建模竞赛章程》和《全国大学生数学建模竞赛参赛规则》（2019 年修订稿，以下简称为“竞赛章程和参赛规则”，可从全国大学生数学建模竞赛网站下载）。

我们完全清楚，在竞赛开始后参赛队员不能以任何方式，包括电话、电子邮件、“贴吧”、QQ 群、微信群等，与队外的任何人（包括指导教师）交流、讨论与赛题有关的问题；无论主动参与讨论还是被动接收讨论信息都是严重违反竞赛纪律的行为。

我们完全清楚，抄袭别人的成果是违反竞赛章程和参赛规则的行为；如果引用别人的成果或资料（包括网上资料），必须按照规定的参考文献的表述方式列出，并在正文引用处予以标注。

我们以中国大学生名誉和诚信郑重承诺，严格遵守竞赛章程和参赛规则，以保证竞赛的公正、公平性。如有违反竞赛章程和参赛规则的行为，我们将受到严肃处理。

我们授权全国大学生数学建模竞赛组委会，可将我们的论文以任何形式进行公开展示（包括进行网上公示，在书籍、期刊和其他媒体进行正式或非正式发表等）。

我们参赛选择的题号（从 A/B/C/D/E 中选择一项填写）：A

我们的报名参赛队号（12 位数字全国统一编号）：202010038016

参赛学校（完整的学校全称，不含院系名）：南京大学

参赛队员 (打印并签名)：1. 陈万驰 181840033 数学系

2. 姜玉骅 181870080 工程管理学院

3. 佘帅杰 181860077 计算机科学与技术系

指导教师或指导教师组负责人 (打印并签名)：教练组

（指导教师签名意味着对参赛队的行为和论文的真实性负责）

日期：2020 年 08 月 1 日

（请勿改动此页内容和格式。此承诺书打印签名后作为纸质论文的封面，注意电子版论文中不得出现此页。以上内容请仔细核对，如填写错误，论文可能被取消评奖资格。）

赛区评阅编号（由赛区组委会填写）：

2020 高教社杯全国大学生数学建模竞赛

编 号 专 用 页

赛区评阅记录（可供赛区评阅时使用）：

评阅人						
备注						

送全国评阅统一编号（由赛区组委会填写）：

全国评阅随机编号（由全国组委会填写）：

（请勿改动此页内容和格式。此编号专用页仅供赛区和全国评阅使用，参赛队打印后装订到纸质论文的第二页上。注意电子版论文中不得出现此页。）

穿越沙漠

摘要

关键字： 新冠肺炎疫情　SEIR 模型　混样检测

一、 问题重述与分析

考虑如下的小游戏：玩家凭借一张地图，利用初始资金购买一定数量的水和食物（包括食品和其他日常用品），从起点出发，按照游戏规则在沙漠中行走。途中会遇到不同的天气，也可在矿山、村庄补充资金或资源，目标是在规定时间内到达终点，并保留尽可能多的资金。

1. 假设只有一名玩家，在整个游戏时段内每天天气状况事先全部已知。对于“第一关”和“第二关”地图，由于天气已知，所以玩家在游戏过程中应当始终向着某个特殊点（如终点、村庄、矿山）行走，而不会在中途因随机天气原因改变行程方向。所以将两张地图进行简化，只保留四种特殊点，并计算出各个点之间在一般情况下的最短距离（即不考虑沙暴天气，晴朗和高温天气都行走）。由此可以通过蒙特卡洛模拟计算出玩家的最佳策略。

2. 假设只有一名玩家，玩家仅知道当天的天气状况，可据此决定当天的行动方案。由于天气随机，所以需要讨论不同天气对玩家策略的影响。另外，是否位于矿山又是决定玩家决策的一个重要因素。所以我们基于这些因素给出玩家策略。利用附件中的“第一关”和“第二关”的天气数据，我们首先对天气概率进行估计，发现在沙漠中晴朗和高温天气占绝大部分时间，沙暴天气占少部分时间，由此可以以频率代替概率，得到三种天气出现的概率。然后利用随机生成数得到一组随机天气序列。在问题一的分析与模型的基础上，我们对给定的随机天气进行模拟，并最终得到玩家在“第三关”和“第四关”的合适的策略。

3. 现有名玩家，他们有相同的初始资金，且同时从起点出发。若某天其中的任意名玩家均从区域 A 行走到区域 B，则他们中的任一位消耗的资源数量均为基础消耗量的倍；若某天其中的任意名玩家在同一矿山挖矿，则他们中的任一位消耗的资源数量均为基础消耗量的倍，且每名玩家一天可通过挖矿获得的资金是基础收益的；若某天其中的任意名玩家在同一村庄购买资源，每箱价格均为基准价格的倍。其他情况下消耗资源数量与资源价格与单人游戏相同。

（1）假设在整个游戏时段内每天天气状况事先全部已知，每名玩家的行动方案需在第天确定且此后不能更改。对第五关地图进行简化与分析后，我们发现这一关的最佳路径是直接从起点出发前往终点，不前往矿山。从而该问题转化为二人选择完全相同的对称博弈问题。为解决此类问题，可以采用混合策略方法。因为游戏中第二天是高温天气，玩家可选择是否停留，由此有两种选择。利用蒙特卡洛模拟来预测玩家在两种选择下的四种路径的平均收益，得到满足 Nash 均衡的混合策略。

（2）假设所有玩家仅知道当天的天气状况，从第天起，每名玩家在当天行动结束后

均知道其余玩家当天的行动方案和剩余的资源数量，随后确定各自第二天的行动方案。

二、 模型假设

1. 假设人群中所有个体都有被感染的概率。
2. 假设被感染个体痊愈后，会产生抗体，不会再被感染。
3. 假设所有感染者是同质的，即病情的严重程度、死亡率相同。
4. 鉴于 1 月 23 日武汉市封城，假设武汉市总人口的不变。
5. 假设核酸检测准确率为 100%。
6. 假设混样检测的结果不会因混样而改变，即若至少有一人感染，混样检测呈阳性；若无人感染，混样检测呈阴性。

三、 符号说明

表 **1** 符号说明表

参数	定义	单位
$Weight$	负重上限	千克
Q	初始资金	元
t	天数	天
T	总天数	天
m_w	每箱水的质量	千克/箱
m_f	每箱食物的质量	千克/箱
p_w	水的基准价格	元/箱
p_f	食物的基准价格	元/箱
n_{sw}	晴朗天气下水的基准消耗量	箱
n_{hw}	高温天气下水的基准消耗量	箱
n_{ow}	沙暴天气下水的基准消耗量	箱
n_{sf}	晴朗天气下食物的基准消耗量	箱
n_{hf}	高温天气下食物的基准消耗量	箱
n_{of}	沙暴天气下食物的基准消耗量	箱
n	玩家数	人
P_t	第 t 天开始时玩家所处的位置	/
W_t	第 t 天开始时玩家剩余的水	箱
F_t	第 t 天开始时玩家剩余的食物	箱
Q_t	第 t 天开始时玩家剩余的资金	元
S_t	第 t 天玩家所处的地点特征	/
Wea_t	第 t 天的天气	/
Q_{Mine}	基础收益	元

四、 模型建立与求解

4.1 游戏模型的建立

我们首先将该游戏利用数学语言加以描述。显然该局游戏的最终目的是使玩家到达终点时的收益最大，即

$$\max Q_{30}+\frac{1}{2}p_wW_{30}+\frac{1}{2}p_fF_{30} \tag{1}$$

其中 Q_t,W_t,F_t 分别表示第 t 天时玩家所剩下的资金、水和食物量。如果玩家在第 30 天前到达终点，则其各个属性将会在未来几天视作不变，所以我们以第三十天为统一结束时间。该目标函数有如下约束：

$$Q_t=Q_{t-1}+Q_{Mine}Mine_t-Shop_t[2p_fShopF_t+2p_wShopW_t] \tag{2}$$

其中

$$Mine_t=\begin{cases}0, & \text{如果第 t 天不挖矿}\\ 1, & \text{如果第 t 天挖矿}\end{cases}$$

$$Shop_t=\begin{cases}0, & \text{如果第 t 天不购物}\\ 1, & \text{如果第 t 天购物}\end{cases}$$

即每天结束时的资金等于前一天的资金加上当天挖矿获得的 1000 元（如果挖矿的话），再减去在村庄购买食物和水花费的钱（如果购买的话）。其中 $ShopF_t$ 和 $ShopW_t$ 分别表示玩家在第 t 天购买的食物量和水量（如果购买的话）。

$$F_t=F_{t-1}-2Move_t\triangle F_t-3Mine_tMove_t\triangle F_t-(1-Move_t-Mine_t)\triangle F_t+Shop_tShopF_t \tag{3}$$

即每天结束时的食物量等于前一天的食物量减去当天的食物消耗量再加上在村庄购买的食物量（如果购买的话）。

$$W_t=W_{t-1}-2Move_t\triangle W_t-3Mine_t\triangle W_t-(1-Move_t-Mine_t)\triangle W_t+Shop_tShopW_t \tag{4}$$

即每天结束时的水量等于前一天的水量减去当天的水消耗量再加上在村庄购买的水量（如果购买的话）。

$$P_t = P_{t-1}(1 - Move_t) + \bar{P}_{t-1}Move_t \tag{5}$$

其中 P_t 表示第 t 天的位置，$\bar{P}_t$ 表示第 t+1 天可以到达的几个相邻区域之一，

$$Move_t = \begin{cases} 1, & \text{如果第 t 天移动} \\ 0, & \text{如果第 t 天不移动（包括挖矿、停留）} \end{cases}$$

$$Shop_t \leqslant If_0[(P_t - C_1)(P_t - C_2)\cdots(P_t - C_n)] \tag{6}$$

$$Mine_t \leqslant If_0[(P_{t-1} - K_1)(P_{t-1} - K_2)\cdots(P_{t-1} - K_m)] \tag{7}$$

其中 $C_1, C_2, \cdots, C_n$ 表示 n 个村庄的位置，$K_1, K_2, \cdots, K_m$ 表示 m 个矿山的位置，

$$If_0(x) = \begin{cases} 1, & if \quad x = 0 \\ \\ 0, & if \quad x \neq 0 \end{cases}$$

$\triangle F$ 和 $\triangle W$ 分别表示第 t 天基础消耗的食物量和水量，由 Lagrange 插值公式：

$$L(x) = \sum_{i=1}^{n} f(x_i)\frac{\Pi_{j\neq i}(x - x_j)}{\Pi_{j\neq i}(x_i - x_j)}$$

可得

$$\triangle F_t = n_{sf}\frac{Wea_t^2 - 3Wea_t + 2}{2} + n_{hf}\frac{Wea_t^2 - 2Wea_t}{-1} + n_{of}\frac{Wea_t^2 - Wea_t}{2} \tag{8}$$

$$\triangle W_t = n_{sw}\frac{Wea_t^2 - 3Wea_t + 2}{2} + n_{hw}\frac{Wea_t^2 - 2Wea_t}{-1} + n_{ow}\frac{Wea_t^2 - Wea_t}{2} \tag{9}$$

其中

$$Wea_t = \begin{cases} 0, & \text{晴朗天气} \\ 1, & \text{高温天气} \\ 2, & \text{沙暴天气} \end{cases}$$

表示第 t 天的天气情况。

4.2 问题一的分析与求解

4.2.1 地图简化

我们首先引入 Dijkstra 算法，这是一种在有向赋权图中求两点之间最短路径的高效算法。

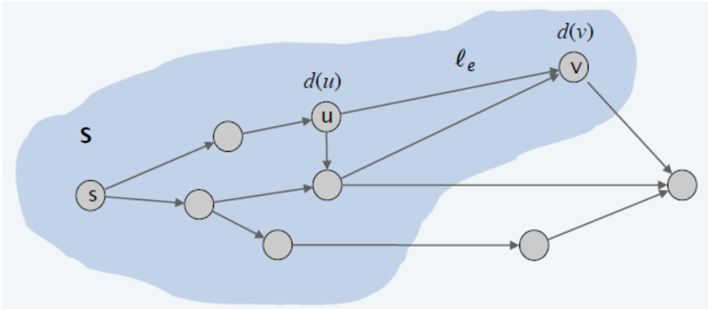


图 1　**Dijkstra** 算法示意图

如图1, 该图中 s 是起点，最右侧的点是终点。$d(u)$ 表示从点 s 到点 u 的最短距离。之后执行如下操作：

(1) 初始化 $S = s, d(s) = 0$;

(2) 反复寻找未探索过的点 v，使得 min　$\pi(v)$，将 v 添加到 S 中，并且 $d(v) = \pi(v)$，其中

$$\pi(v) = \min_{e=(u,v):u\in S} d(u) + distance(arg \min_{e=(u,v):u\in S} d(u), v)$$

最终当终点属于 S 时，可以知道起点到终点的最短路径长度，从而问题求解。

可以看出，问题一的核心关键在于玩家是否前往矿山挖矿、在矿山连续挖几天矿、在挖矿之后是否需要前往村庄补给物资以及是否可以在矿山和村庄之间往返。我们将问题作如下简化：

选取图中起点、终点、村庄和矿山作为特殊点，利用 Dijkstra 算法计算出每两个特殊点之间的最短路径（即不考虑沙暴天气的影响下，所需最少的天数），如图2所示。

该图是一个完全图（即图中每两点之间都有一条边），且每条边上的数字代表着两个特殊点之间的最短路径（受天气影响可能会有变化）。对于问题一，因为天气是预先给定的，整局游戏没有随机因素，所以我们假定玩家在游戏途中始终向着某个特殊点（终点、村庄、矿山）前进并且选择最合适的路径，而非漫无目的地随机移动，即游戏旅程由几条有向线段叠加而成：

$$\overrightarrow{P} = \sum_{k=1}^{n} \overrightarrow{P_{i_k} P_{i_{k+1}}} \tag{10}$$

其中 P_0, P_1, P_2, P_3 分别表示起点、村庄、矿山和终点，$i_k \leqslant 3, \quad k = 1, 2, \ldots, n.$

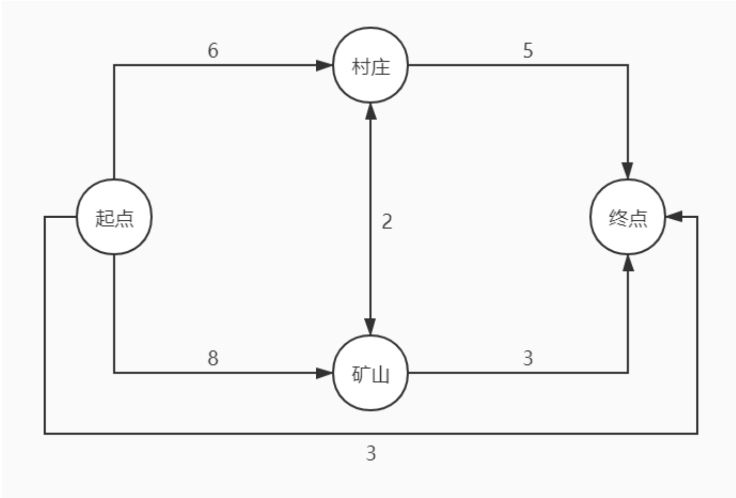


图 **2**　简化后的 **Map1**

因此，问题一可简化为玩家在这几个特殊点之间的运动问题。由于第一关地图中村庄在由起点去往矿山的必经之路上，所以应该考虑在玩家第一次到达村庄时适量补充水和食物，然后前往矿山尽可能多地挖矿，在水和食物只够支撑玩家返回村庄时停止挖矿，选择返回村庄。在村庄补给一定量的水和食物（可以通过计算之后几天返回终点时的行程来确定需要购买的水和食物的量，以使得玩家到达终点时水和食物剩余量都为 0，从而杜绝资产损失。）

4.2.2 路径方案

我们将针对问题一的策略方案用流程图的形式展示出来, 如图3。根据这样的流程

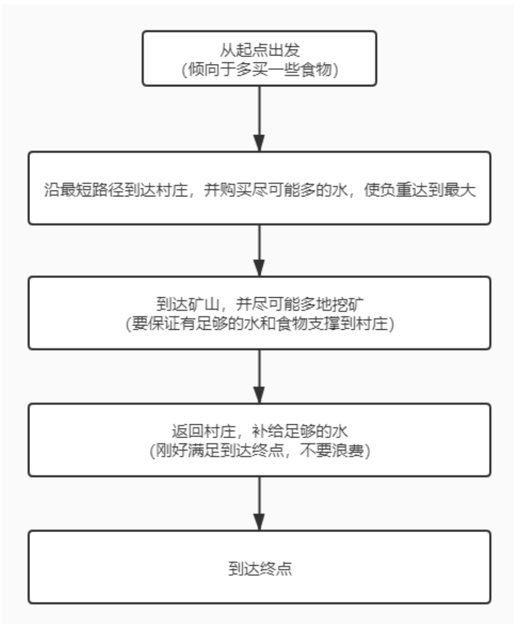


图 **3**　第一关策略方案流程图

图，我们进行问题分析和代码求解。玩家在游戏过程中的可选择性因素有：起点处的物资储备、在村庄的物资购买、在矿山挖矿的持续天数、返回村庄时的物资购买等。对此，我们对起点处的物资准备和在矿山挖矿的持续天数进行有范围穷举，对两次村庄物资购买分别采用精准计算的方法，借助 Python 编程语言求解。

其中，第一次到达村庄时购买物资的策略为：只购买水，且尽量让负重达到最大。即

$$W_{buy1} = \min\{[\frac{Weight - m_w W_{c1} - m_f F_{c1}}{m_w}], \frac{Q_{c1}}{p_w}\} \tag{11}$$

$Weight$ 为最大负重，W_{c1} 为第一次到达村庄时剩余的水量，F_{c1} 为第一次到达村庄时剩余的食物量，m_w 和 m_f 分别表示一箱水和一箱食物的质量，Q_{c1} 为第一次到达村庄时的剩余资金，p_w 为一箱水的价格。

第二次到达村庄时购买物资的策略为：准确满足之后到达终点所需的全部水和食物，使得到达终点时不剩余任何水和食物。具体代码实现参见附件。

经过代码验证，最终找到最佳的策略方案，路径如图4所示。

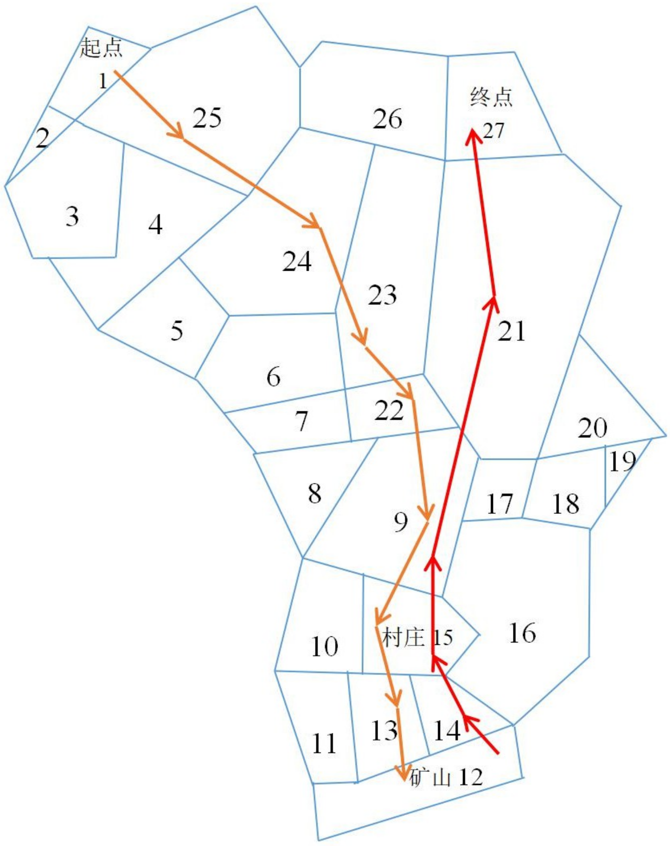


图 4　第一关路径示意图

具体的游戏路径：经历 8 天时间到达村庄 15（其中有 2 天沙暴天气，只能停在原

地），当天在村庄购买水 163 箱。第二天立即前往矿山，第 10 天到达矿山，第 11-17 天进行挖矿，第 18 天停留在矿山。第 19-20 天前往村庄，并在第 20 天当天购买水 36 箱、食物 40 箱。第 21-23 天前往终点，最终于第 23 天到达终点，水和食物全部消耗完毕。最终剩余资金 10430 元。完整路径参见附件。

五、 问题二

此时玩家仅仅知道当前的天气状况，可据此决定当天的行动方案，即行走或停留或挖矿，整体上的策略仍然分为两种，直接行走到终点或先到矿山挖矿再到终点。玩家可以根据当前天气状况，获取各个行动方案的未来收益，作以下讨论。

1. 若玩家不在矿山：此时玩家需要决策行走或停留。
   (a) 沙暴天气必然停留。题目强制要求。
   (b) 晴朗天气必然行走。晴朗天气行走消耗资源最少，若晴朗天气不行走，则不存在能够行走的天气状况。
   (c) 高温天气需要根据地图参数条件决策。当前时刻为高温，假设下一时刻为晴朗。决策为行走的物资消耗为，水：$2n_{hw}$，食物：$2n_{hf}$。决策为停留等下一时刻晴天再行走的物资消耗为，水：$n_{hw}+2n_{sw}$，食物：$n_{hf}+2n_{sf}$。即满足：

$$\begin{cases} 2n_{sw} < n_{hw} \\ 2n_{sf} < n_{hf} \end{cases} \tag{12}$$

   当前时刻停留的物资消耗更少，因此在时间宽裕的条件下可以考虑在高温天气停留一天。如果当前时刻为高温天气，下一时刻仍为高温天气，另做讨论。

2. 若玩家在矿山：此时玩家需要决策行走或停留或挖矿
   (a) 晴朗天气必然挖矿。晴朗天气挖矿消耗资源最少，若晴朗天气不挖矿，则不存在应挖矿的天气
   (b) 高温条件下需要根据地图参数条件决策。决策为挖矿的收益为：$Q_{mine}-3(n_{hw}p_w+n_{hf}p_f)$，决策为停留的收益为：$-(n_{hw}p_w+n_{hf}p_f)$。即满足下列条件：

$$Q_{mine} > 2(n_{hw}p_w+n_{hf}p_f) \tag{13}$$

   则挖矿收益更高，因此决策为挖矿，否则考虑停留行走。其中 Q_mine 为挖矿基本收益，n_{hw}、nhf 为高温下水、食物的基本消耗量。以上为考虑当前资源剩余量，若资源剩余量充足，考虑挖矿，若不足，则考虑行走。
   (c) 沙暴天气需要根据地图参数条件决策。沙暴天气只能挖矿或停留。决策为挖矿的收益为：$Q_{mine}-3(n_{ow}p_w+n_{of}p_f)$，决策为停留的收益为：$-(n_{ow}p_w+n_{of}p_f)$。若

满足下列条件：

$$Q_{mine} < 2(n_{ow}p_w + n_{of}p_f) \tag{14}$$

则挖矿收益，因此决策为挖矿，否则停留。其中 n_{ow}、nof 为沙暴天气下水、食物的基本消耗量。

由于天气状况未知，因此本题通过蒙特卡罗模拟的方式，寻求玩家动态最佳决策的期望值。实际模拟中，假设每天天气状况相互独立，天气状况可根据第一关、第二关的天气的先验概率分布随机生成，即按 2:3:1 的概率随机生成晴朗、高温、沙暴三种天气状况。若地图条件满足式 (12)，即当天为高温天气且下一天为晴朗时，当天选择停留消耗的资源更少。由于存在连续几天出现高温天气的可能性，所以高温天气选择停留可能会产生恶性循环。因此，我们以概率形式表示我们的决策，0.4 的概率 (下一天为晴天) 选择停留，0.6 的概率选择行走。

下面就三、四关作具体讨论：

5.1 第三关

第三关的地图不存在村庄，不存在沙暴天气。因此模型简化了许多。此时玩家路线仅需考虑是直接去终点，还是先经过矿山挖矿再到终点，如图 5所示。

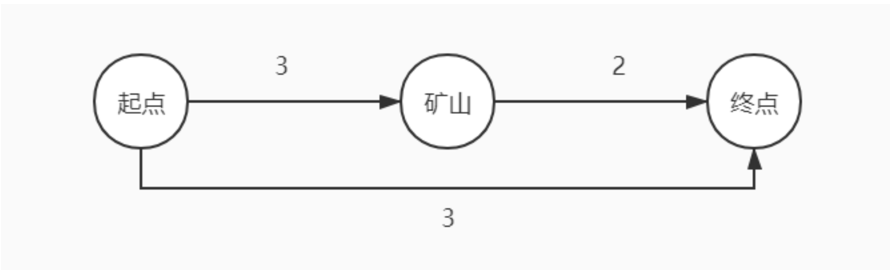


图 **5** 简化后的 **Map3**

第三关地图满足式 (12)， (13)，不满足 (14)，即玩家的策略如下：

1. 晴朗天气：在矿山则挖矿；不在矿山则行走
2. 高温天气：在矿山不挖矿，0.4 概率停留，0.6 概率行走；不在矿山 0.4 概率停留，0.6 概率行走。

天气情况按 2:3 随机生成晴朗、高温。用蒙特卡罗分别对玩家直奔终点和先去矿山再去终点两种路线模拟 1000 轮，每次模拟 100 次，取最终资金最高的一次作为本轮最终资金。结果见图 6

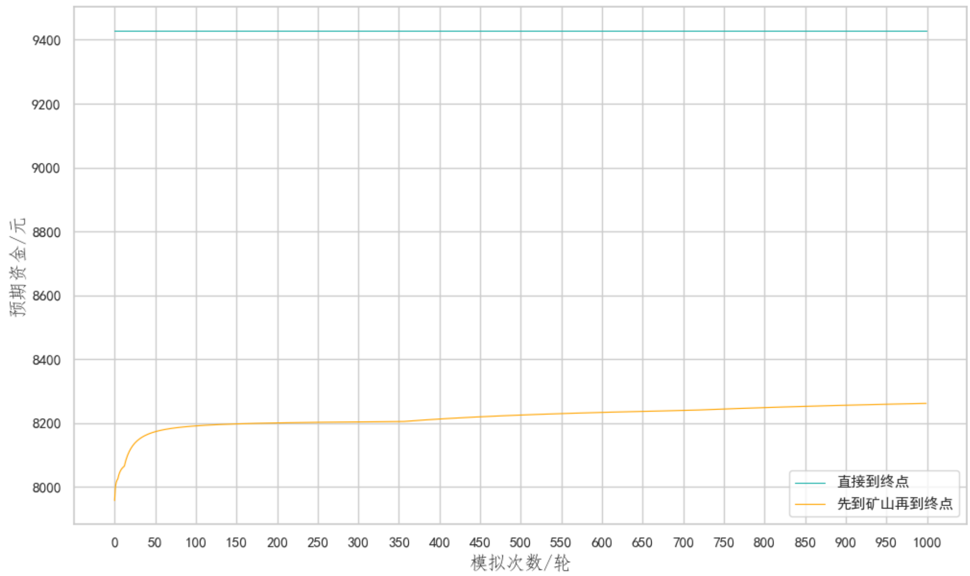


图 **6**　第三关蒙特卡罗模拟结果

可见直接到终点的最佳预期资金稳定在 9450 元左右，先去矿山再去终点的预期最佳资金稳定在 8100 元上下。故第三关地图玩家的路线决策应为直奔终点。

5.2 第四关

第四关的地图有一个村庄、一个矿山且存在沙暴天气。模型较第三关复杂。由于起点到终点的距离等于起点到矿山的距离加上矿山到终点的距离，故玩家必然会挖矿。简化图如图 7

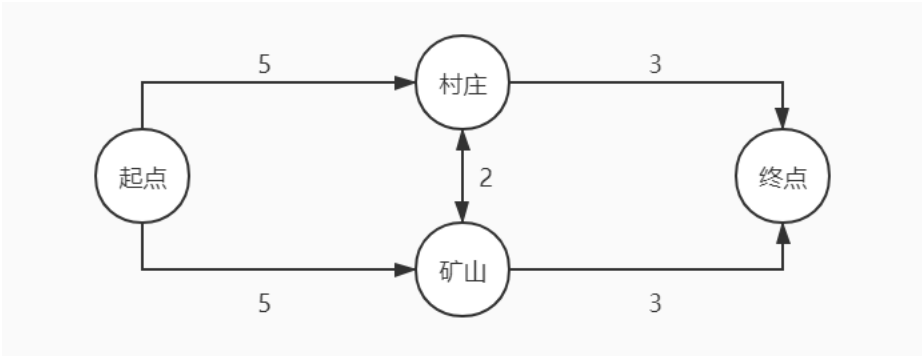


图 **7**　简化后的 **Map4**

第四关地图不满足式 (12)， (13)，满足式 (14)。即玩家的策略如下：

1. 晴朗天气：在矿山则挖矿；不在矿山则行走。
2. 高温天气：在矿山则挖矿；不在矿山则以 0.4 概率停留，0.6 概率行走。

3. 沙暴天气：在矿山则挖矿；不在矿山只能停留。

天气情况按 2:3:1 随机生成晴朗、高温和沙暴。用蒙特卡罗对玩家路线随机模拟 1000 轮，每轮模拟 500 次，取最终资金的最高的一次作为本轮的资金。结果见图 8

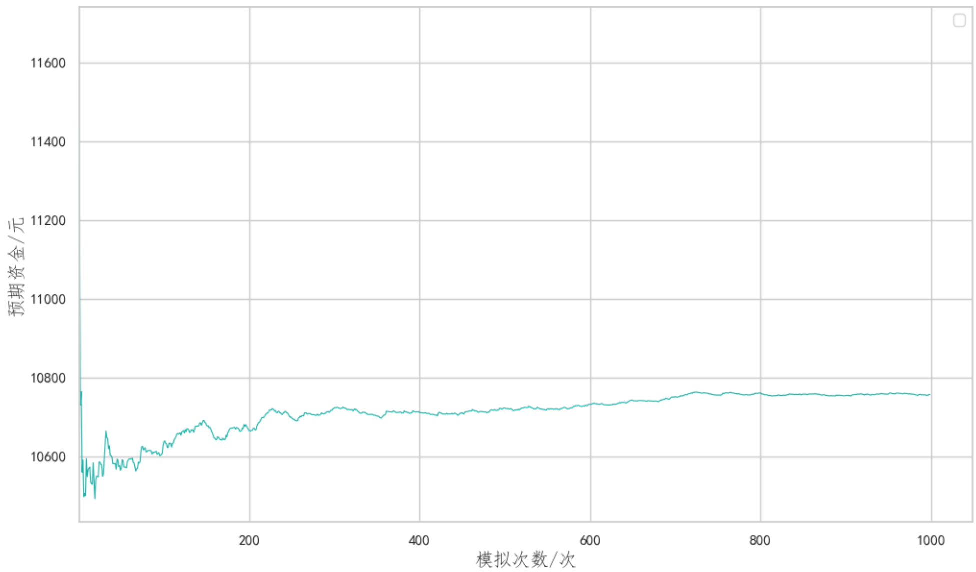


图 **8**　第四关蒙特卡罗模拟结果

可见由于天气带来的随机性，最佳预期资金平均值趋近于 10800。现对于一个表现较好的蒙特卡罗模拟例子作分析。随机生成的天气为 [晴朗, 高温, 高温, 高温, 晴朗, 晴朗, 高温, 高温, 高温, 晴朗, 高温, 晴朗, 高温, 高温, 晴朗, 沙暴, 沙暴, 高温, 沙暴, 高温, 沙暴, 沙暴, 高温, 沙暴, 高温, 晴朗, 沙暴, 晴朗, 高温, 高温]。该天气情况为事实生成的，玩家仍然仅知道当天的天气情况。针对该天气状况，玩家的最优动态策略路径：起点购买 240 箱水、240 箱食物，第 5 天晚上到达矿山 (其中经历沙暴延迟一天)，挖矿 6 天，第 13 天早上离开，并于 16 号早上到达终点，最终资金为 12590。玩家详细路径如图：

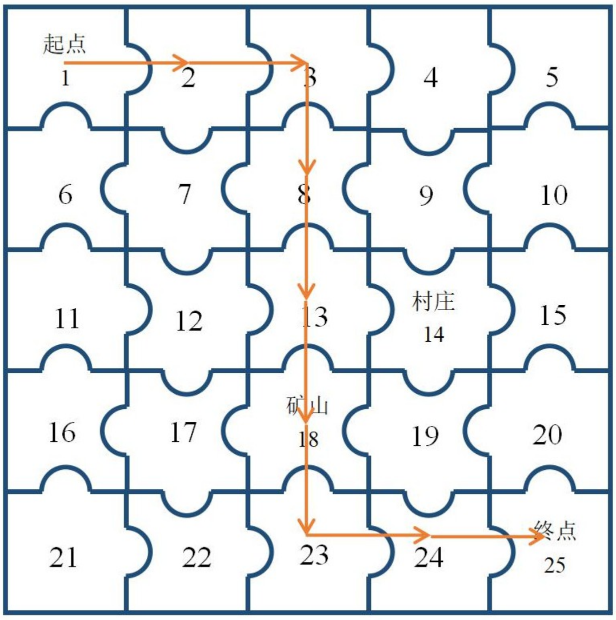


图 **9**　一个最优决策路径的图例

由于天气因素的随机性以及玩家动态决策的随机性会导致玩家的最佳路径不同、最终资金不同。上述蒙特卡罗例子只是某种特定情况下的，最佳路径，不具有典型代表性。该例子中，玩家不需要经过村庄是因为前期行走过程中天气以晴朗为主，消耗物资较少。若在其它天气情况下，可能需要经过村庄补给。

六、 问题三

6.1 第五关

第五关的地图较为简单，可简化为如图图 10所示，我们先从单人游戏开始分析。对该图分析可以发现，本关有两种路径：一是直接由起点奔向终点，一是从起点先到矿山，再到终点。第一种路径的最少花费是 $(6\times2+18)\times5+(8\times2+18)\times10 = 150+340 = 490$。因为天数较少，路径结构较为简单，经过简单计算后发现在第五关地图中，最佳路径是从起点直奔终点，不经过矿山，这样只需实际行走三天。

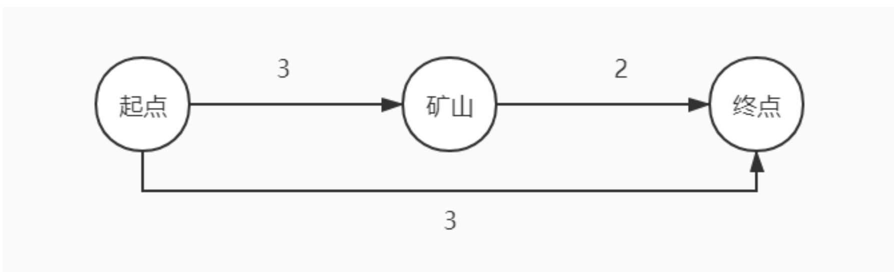
。

图 **10**　简化后的 **Map5**

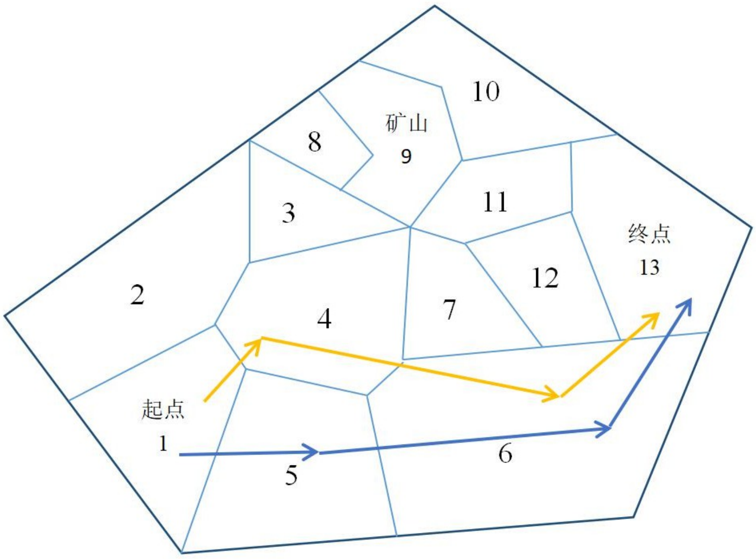


图 **11** **Map5** 的两种最优路径

对于二人游戏，由于两名玩家都可以认为是“绝对自私，完全理智”的，所以可以采用静态博弈理论，即两个决策者都知道所有信息（即共同知识），两人同时做出决策，这样的博弈称为完全信息的静态博弈。在该博弈模型中，参与人是两名玩家，每个人的策略空间有两个元素（即是否在高温天停留），每个人的效用函数即是自己最终剩余资金的期望值。

游戏的博弈中参与人集合可以用 $N = 1, 2$ 表示，1 表示玩家 1，2 表示玩家 2。玩家 1 的可能的策略记作 $a \in A = 1, 2$，1、2 分别表示选择第 2 天不停留和停留；玩家 2 的可能的策略记作 $b \in B = 1, 2$，1、2 分别表示选择第 2 天不停留和停留。A 和 B 分别是参与人 1 和 2 的策略空间。

对于两名玩家的每一个策略组合 (a, b)，Map5 中都有两种路径，且完全等价，所以我们用计算机进行蒙特卡洛模拟，来计算在该种策略组合下两名玩家的平均收益。用 $P(a, b)$ 和 $Q(a, b)$ 分别表示玩家 1 和玩家 2 一次游戏的期望收益，这可以作为玩家 1 和玩家 2 的效用函数。经计算得

$$P = (p_{ij})_{2\times2} = \begin{pmatrix} 8255.09 & 8419.88 \\ 8567.58 & 8402.51 \end{pmatrix} \tag{15}$$

$$Q = (q_{ij})_{2\times2} = \begin{pmatrix} 8255.09 & 8567.44 \\ 8420.16 & 8402.51 \end{pmatrix} \tag{16}$$

可以近似认为

$$P = Q'$$

设玩家 1 采取策略 i 的概率为 $p_i(i=1,2)$，玩家 2 采取策略 j 的概率为 $q_j(j=1,2)$，记行向量 $\mathbf{p}=(p_1,p_2),\mathbf{q}=(q_1,q_2)$，满足

$$0\leqslant p_i\leqslant 1,\quad \sum_{i=1}^{2}p_i=1,\quad 0\leqslant q_j\leqslant 1,\quad \sum_{j=1}^{2}q_j=1$$

的概率向量分别构成玩家 1 和玩家 2 的混合策略空间。

按照混合策略下的 Nash 均衡理论，计算得 $p_1=q_1=0.346,\quad p_2=q_2=0.654$。也即玩家 1 和玩家 2 应该以 34.6% 的概率选择不停留，以 65.4% 的概率选择停留。由于两名玩家是完全对称的，所以认为两人所采取的策略应当是完全相同的。我们用蒙特卡洛进行大量模拟两人的游戏过程，通过改变两人选择不停留的概率（两人始终保持概率相同），步长为 0.001，最终得到个人的期望收益随概率的变化如图12。

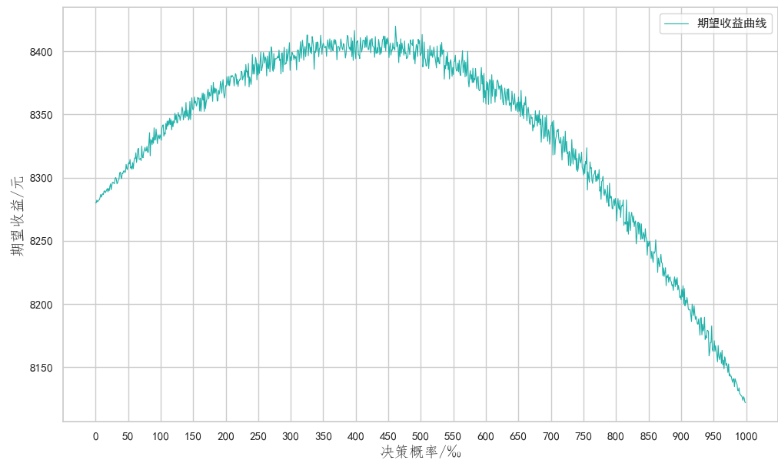


图 **12**　个人期望收益随停留概率变化曲线

从图中可以看出，个人期望收益在概率接近 0.35 到达峰值，这也与我们计算的 0.346 概率高度吻合，所以可以说明通过混合策略的 Nash 均衡理论计算出的结果是较为准确可靠的。

七、 灵敏度分析

在第三关、第四关模拟中，由于未知天气，原模型中假设晴朗、高温、沙暴天气出现比例为 2:3:1，根据这一比例确定了第三关、第四关预期的最终资金。下面我们来分析晴朗、高温、沙暴天气出现比例对第三关、第四关最终资金的灵敏度。

7.1 第三关

我们设置 5 组晴朗、高温天气比例，E1：0.4:0.6；E2：0.45:0.55；E3: 0.5:0.5；E4：0.35:0.65；E5：0.3:0.7 分别用上述天气比例蒙特卡罗模拟第四关 100 轮，每轮 100 次，将每轮最终资金的最大值作为该轮的最终资金。分别得到从起点直接去终点的预期最终资金、从起点先到矿山再到终点的预期最终资金，结果见图 13、图 14。

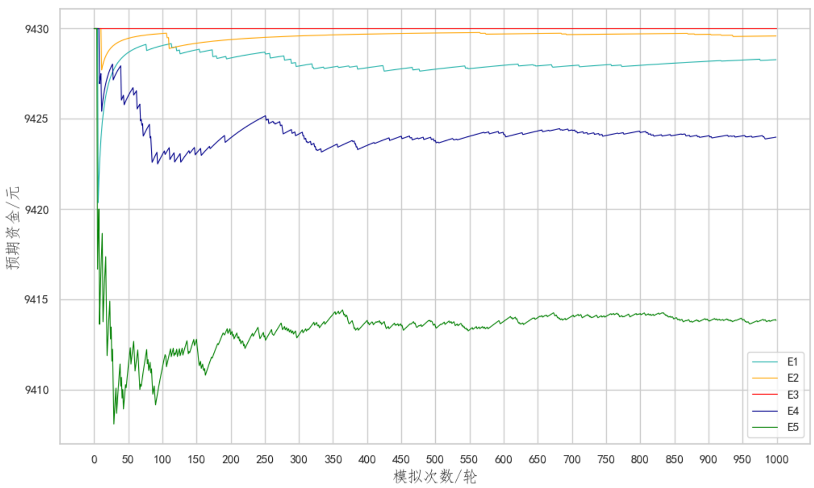


图 **13**　第三关：起点到终点预期最终资金

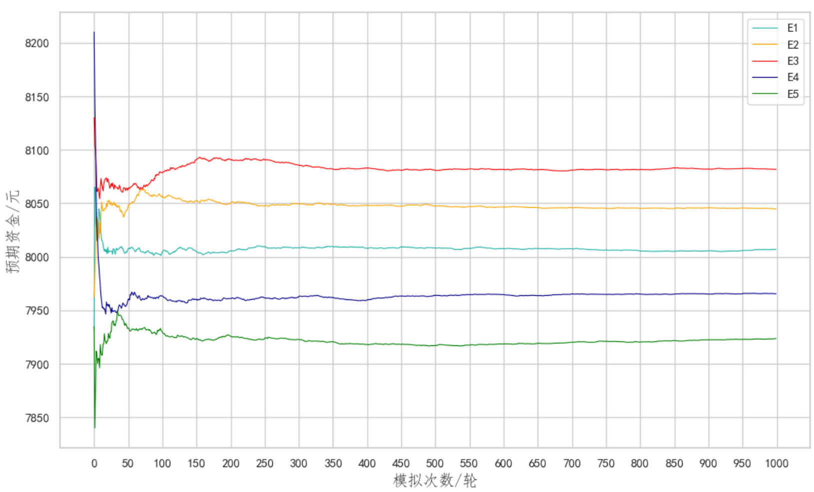


图 **14**　第三关：起点到矿山再到终点预期最终资金

观察发现，晴朗天气比例的降低或高温天气比例的增高，会导致最终资金的减少。这也正确反映了天气情况对玩家策略、资金的影响。但总的来说，仍然是从起点直接到终点的最终资金比从起点到终点再到矿山的高，验证了我们第三关的结果的正确性。

7.2 第四关

我们设置 5 组晴朗、高温、沙暴天气比例，E1：10:15:5；E2：8:17:5；E3: 12:13:5；E4：11:15:4；E5：11:16:3 分别用上述天气比例蒙特卡罗模拟第四关 500 轮，每轮 500 次，将每轮最终资金的最大值作为该轮的最终资金，结果见图 15：

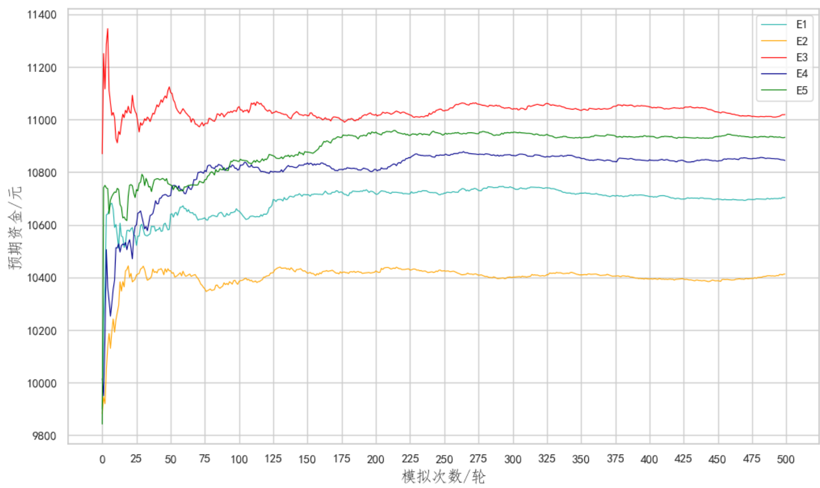


图 **15**　第四关预期最终资金

观察发现，晴朗天气比例越高、高温和沙暴天气比例越低，会促使最终资金的增加，可见天气随机性对结果会有不小的影响，所以第四关玩家只知道当天的天气，意味着全局最优策略和最高资金也是随机的，与天气情况的随机性息息相关。这样的结果与我们的设想符合。

7.3 第六关

第六关是本题中难度最大的一关，不太可能使用前述算法将可能的情况遍历或模拟，我们用动态评估的方法给出在具体情况下的求解方法。

首先，我们在游戏开始之前，沿用前述方法对天气进行随机模拟，得到天气序列。然后我们使用计算机模拟玩家进行游戏。当玩家走到某一区域时，搜索与该区域相邻的区域作为下一天可采取的行动空间（包括购买物资、采矿等）。之后我们利用蒙特卡洛树搜索方法对这几个区域进行大量模拟，得出几个区域最终的期望收益。即

$$\bar{Q} = \frac{\sum_{i=1}^{n} Q_i}{n} \tag{17}$$

其中 $\bar{Q}$ 表示某一区域的期望收益，Q_i 表示第 i 次模拟的最终收益。由大数定律可知，当模拟次数足够大时，该区域的期望收益将会收敛于一个值，我们认为该值能代表该区域能给玩家带来的平均收益。所以玩家应在几个可选区域中选取期望收益最高的一个并在下一天执行。

我们以区域 17 为例来演示这一过程。如图16所示，区域 17 周围有四个区域，分别是 12、16、18、22，红色数字是各个区域经过大量模拟得到的期望收益，其中 16 位玩家已走过的区域，12 为另外两位玩家（在计算机模拟中固定他们的运动轨迹）要经过的区域，所以这两个区域的期望收益较低；而 18 是矿场，玩家可以在矿场开采矿石，所以期望收益较高，这也符合实际。最终玩家选择期望收益最高的 18 区域并在第二天前往到达。

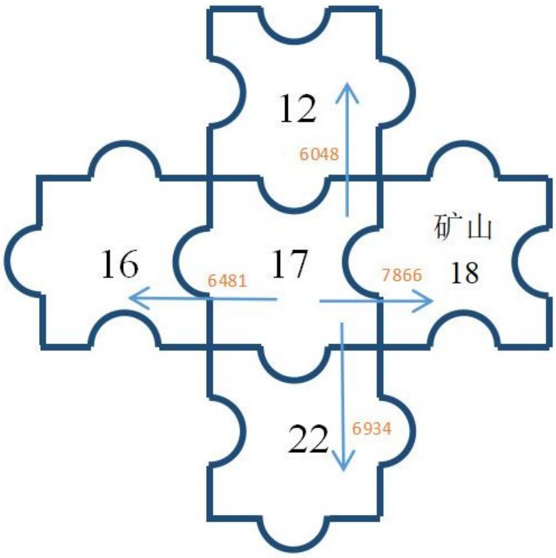


图 **16**　区域 **17** 期望收益计算过程示意图

如图17所示，我们事先给定两个机器人（橙色和蓝色）的行动方案，玩家（红色）在每一次行动过后得知机器人的行动方案，并基于这些信息以及前述的期望收益评估方法来规划第二天的行程。最终，玩家路径几乎完美的与其他两个机器人的路径错开。

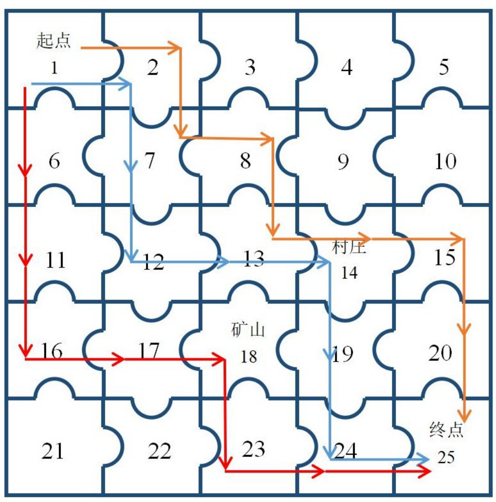


图 **17**　动态评估给出的玩家最优路径

参考文献

参考文献