关于“施了魔法的羊”的一些想法

因为没有多少相关知识，才疏学浅，所以我只好用我比较熟悉的方式来切入这个问题试试看......

1 模型与算法

因为管理者的决策只和当前时刻有多少只黑羊，有多少只白羊有关，所以我尝试着用概率结合上动态规划来研究这个问题。既然决策只和黑白羊各自的个数有关，我用 (a, b) 表示一种有 a 只黑羊，b 只白羊的状态。记 $f(a, b)$ 表示有 a 只黑羊和 b 只白羊的情况，在使用最优的策略下，最终黑羊的期望个数。

如果管理者选择在最开始的时候移走若干只白羊（至少一只），设管理者一次操作之后还剩下 c 只白羊 $(0 \leq c < b)$，那么就存在转移

$$f(a, b) \leftarrow \frac{ap}{ap + cq} f(a + 1, c - 1) + \frac{cq}{ap + cq} f(a - 1, c + 1)$$

管理者想要找到最优的策略，就要找到所有转移里面最大的那一个，也就是

$$f(a, b) \leftarrow \max\{a, \max_{0<c<b}\{\frac{ap}{ap + cq} f(a + 1, c - 1) + \frac{cq}{ap + cq} f(a - 1, c + 1)\} \tag{1}$$

这个转移是无后效性的，也就是只要按照 $a + b$ 从小到大计算所有的 $f(a, b)$，那么在算 $f(a, b)$ 的时候，$f(a + 1, c - 1)$ 和 $f(a - 1, c + 1)$ 都已经算好了。所以这个转移可以用递推很快算出来。之所以用了“$\leftarrow$”而不是“$=$”，是因为这个状态转移是不全面的，因为管理者还有可能一只羊也不移走，也就是接下来的这种情况。

如果不移走任何羊，那么在有 a 只黑羊，b 只白羊的情况下。经过足够长的时间，总会有羊发生变化。即在经过足够久之后，有 $\frac{ap}{ap+bq}$ 的概率让一只白羊变黑，有 $\frac{bq}{ap+bq}$ 的概率让一只黑羊变白。对于这种不移走任何羊的情况，可以列出以下转移。

$$f(a, b) \leftarrow \frac{ap}{ap + bq} f(a + 1, b - 1) + \frac{bq}{ap + bq} f(a - 1, b + 1) \quad (a, b > 0) \tag{2}$$

这个方程用到了 $f(a, b), f(a + 1, b - 1), f(a - 1, b + 1)$，它们的两个自变量之和总是常数 a+b，也就是说转移出现了环，所以不能像 (1) 式那样直接递推。所以我只好退而求其次找一个复杂度高一点的算法。仔细想一下，管理者如果不移走任何羊，那么状态可能会从 (a, b)，变成 $(a + 1, b - 1)$，然后又变到 $(a + 2, b - 2)$，接着可能又变到 $(a + 1, b - 1)$ 等等。这样不断变化，最终一定会遇到以下两种情况之一

1. 达到了某个状态 (a', b')，满足 $(a + b = a' + b', a', b' > 0)$，这时候最优策略要求管理者要移走白羊了。也就是说不移走羊的情况结束了。

2. 达到了 $(0, a + b)$ 或者 $(a + b, 0)$。整个过程完全结束了。

所以我们可以直接枚举两个数 $l, r (l < a < r)$，表示 $(l, a + b - l)$ 和 $(r, a + b - r)$ 分别是离 (a, b) 最近的两个会让不移走羊的情况结束的状态。这时候只需要算一下从 (a, b) 开始，不移走任何羊，最终状态变

成 $(l, a+b-l)$ 的概率 p_1 和 $(r, a+b-r)$ 的概率 p_2（当然 $p_1+p_2=1$）。这个用方程组就可以算了，后面再提。因此这部分转移是这样的

$$f(a,b) \leftarrow p_1 \cdot g(l, a+b-l) + p_2 \cdot g(a+b-r, r) \tag{3}$$

其中 $g(a,b)$ 表示有 a 只黑羊和 b 只白羊的情况，刚开始必须先移走至少一只白羊，在最优的策略下，最终黑羊的期望个数。

这时候我们再回来考虑管理者必须要移走白羊的那一部分转移（也就是 (1) 式做的事情），发现利用 $g(a,b)$，(1) 式可以简化成如下形式

$$f(a,b) \leftarrow g(a,b) \tag{4}$$

现在只需要知道 $g(a,b)$ 怎么算就行了。注意到 $g(a,b)$ 直接使用 (1) 式就可以递推出来，写下来也就是

$$g(a,b) = \max\{a, \max_{0<c<b}\{\frac{ap}{ap+cq}g(a+1,c-1) + \frac{cq}{ap+cq}g(a-1,c+1)\} \tag{5}$$

万事俱备，接下来只需要计算出 p_1, p_2 就好了。记 $p(a,b)$ 表示从 (a,b) 开始不移走任何白羊，变成 $(l, a+b-l)$ 的概率（换句话说就是从 (a,b) 开始的 p_1）。注意到

$$p(a,b) = \frac{ap}{ap+bq}p(a+1,b-1) + \frac{bq}{ap+bq}p(a-1,b+1)$$

发现 $p(a,b)$ 之间的转移有环，也不能递推。但是没有关系，我们可以列出如下 $r-l+1$ 个方程

$$\begin{cases} p(l, a+b-l) = 1 \\ \cdots \\ p(a-1,b+1) = \frac{(a-1)p}{(a-1)p+(b+1)q} \cdot p(a,b) + \frac{(b+1)q}{(a-1)p+(b+1)q} \cdot p(a-2,b+2) \\ p(a,b) = \frac{ap}{ap+bq} \cdot p(a+1,b-1) + \frac{bq}{ap+bq} \cdot p(a-1,b+1) \\ p(a+1,b-1) = \frac{(a+1)p}{(a+1)p+(b-1)q} \cdot p(a+2,b-2) + \frac{(b-1)q}{(a+1)p+(b-1)q} \cdot p(a,b) \\ \cdots \\ p(r, a+b-r) = 0 \end{cases}$$

恰好有 $r-l+1$ 个未知数 p，所以用高斯消元等办法解上述方程就可以把需要的 p_1, p_2 算出来了。

至此，运用 (3)(4)(5) 式以及上面这个方程组，就搭建了一个模型，并且有了一个时间复杂度为 $O(n^6)$ 的算法。后来发现 r 只可能取 $a+b$，以及方程组也很有特点，所以复杂度可以降到 $O(n^4)$ 甚至 $O(n^3)$。对于 2000 只黑羊和 2000 只白羊的情况，应该可以运行出结果来。不过如果只是用来观察和验证猜想，就应该没有必要优化速度了。（代码附后）

2 验证

以上是粗略的理论分析，感觉上还没有十足的把握（毕竟我有时候想问题也比较粗心），所以我按照书上的结论验证了一下。书上有说，当不考虑 p, q，即 $p, q = 1$ 的时候

$$\lim_{k\to\infty} f(k,k) - (2k + \frac{\pi}{4} - \sqrt{\pi k}) = 0$$

虽然我看不太懂它的证明，但是拿来用还是很好的。如果 $k = 50$，此时 $2k + \frac{\pi}{4} - \sqrt{\pi k} \approx 88.25$。用上面的算法算了一下，结果是 88.283351，比较准确。我还写了一个在 $p, q = 1$ 的情况下按照书上说的“保证

黑羊总比白羊多”的策略的模拟器，进行了 20000000 次模拟，结果和上面的算法运行出来的结果很吻合。说明这个模型和算法应该没什么问题，至少在 $p, q = 1$ 的时候应该是对的。

3 探究与结论

因为对于 $p, q = 1$，或者说 $p = q$ 的情况，书上已经给出了策略，就是时刻保证黑羊总比白羊多（当然因为知识有限我没有看懂证明...），所以很自然地猜在 p, q 不一定相同的情况下，只需要让产生黑羊的概率比产生白羊的大就好了，也就是保证

$$\frac{ap}{ap+bq} - \frac{bq}{ap+bq} > 0 \qquad (*)$$

我用上面提到的算法和模拟器进行了测试。模拟器的策略是：一旦 $\frac{ap}{ap+bq} - \frac{bq}{ap+bq} \leq 0$，就开始不断移走白羊。

粗试了几组，结果都比较吻合，但是我渐渐发现了有点问题。例如，在有 21 只黑羊，15 只白羊，p=0.5，q=0.8 的时候，动态规划算法给出的结果为：最优策略下黑羊个数的期望值是 28.971949，但是我的模拟器在多次模拟的情况下，结果稳定在 28.964 附近。如果我们默认算法得出的结果是最佳值的话，这就意味着模拟器采用的这种策略可能不是最优的。

事实上，我把算法里面发生的转移打印出来就发现了问题。用表格整理如下（规定 $p = 0.5, q = 0.8$）

-0.30	-1.10	-1.90	-2.70	-3.50	-4.30	-5.10	-5.90	-6.70	-7.50	-8.30	-9.10	-9.90	-10.70	-11.50	-12.30	-13.10	-13.90	-14.70	-15.50	-16.30	-17.10	-17.90	-18.70	-19.50	-20.30	-21.10	-21.90	-22.70	-23.50
0.20	-0.60	-1.40	-2.20	-3.00	-3.80	-4.60	-5.40	-6.20	-7.00	-7.80	-8.60	-9.40	-10.20	-11.00	-11.80	-12.60	-13.40	-14.20	-15.00	-15.80	-16.60	-17.40	-18.20	-19.00	-19.80	-20.60	-21.40	-22.20	-23.00
0.70	-0.10	-0.90	-1.70	-2.50	-3.30	-4.10	-4.90	-5.70	-6.50	-7.30	-8.10	-8.90	-9.70	-10.50	-11.30	-12.10	-12.90	-13.70	-14.50	-15.30	-16.10	-16.90	-17.70	-18.50	-19.30	-20.10	-20.90	-21.70	-22.50
1.20	0.40	-0.40	-1.20	-2.00	-2.80	-3.60	-4.40	-5.20	-6.00	-6.80	-7.60	-8.40	-9.20	-10.00	-10.80	-11.60	-12.40	-13.20	-14.00	-14.80	-15.60	-16.40	-17.20	-18.00	-18.80	-19.60	-20.40	-21.20	-22.00
1.70	0.90	**0.10**	-0.70	-1.50	-2.30	-3.10	-3.90	-4.70	-5.50	-6.30	-7.10	-7.90	-8.70	-9.50	-10.30	-11.10	-11.90	-12.70	-13.50	-14.30	-15.10	-15.90	-16.70	-17.50	-18.30	-19.10	-19.90	-20.70	-21.50
2.20	1.40	0.60	-0.20	-1.00	-1.80	-2.60	-3.40	-4.20	-5.00	-5.80	-6.60	-7.40	-8.20	-9.00	-9.80	-10.60	-11.40	-12.20	-13.00	-13.80	-14.60	-15.40	-16.20	-17.00	-17.80	-18.60	-19.40	-20.20	-21.00
2.70	1.90	1.10	0.30	-0.50	-1.30	-2.10	-2.90	-3.70	-4.50	-5.30	-6.10	-6.90	-7.70	-8.50	-9.30	-10.10	-10.90	-11.70	-12.50	-13.30	-14.10	-14.90	-15.70	-16.50	-17.30	-18.10	-18.90	-19.70	-20.50
3.20	2.40	1.60	0.80	**0.00**	-0.80	-1.60	-2.40	-3.20	-4.00	-4.80	-5.60	-6.40	-7.20	-8.00	-8.80	-9.60	-10.40	-11.20	-12.00	-12.80	-13.60	-14.40	-15.20	-16.00	-16.80	-17.60	-18.40	-19.20	-20.00
3.70	2.90	2.10	1.30	0.50	-0.30	-1.10	-1.90	-2.70	-3.50	-4.30	-5.10	-5.90	-6.70	-7.50	-8.30	-9.10	-9.90	-10.70	-11.50	-12.30	-13.10	-13.90	-14.70	-15.50	-16.30	-17.10	-17.90	-18.70	-19.50
4.20	3.40	2.60	1.80	1.00	0.20	-0.60	-1.40	-2.20	-3.00	-3.80	-4.60	-5.40	-6.20	-7.00	-7.80	-8.60	-9.40	-10.20	-11.00	-11.80	-12.60	-13.40	-14.20	-15.00	-15.80	-16.60	-17.40	-18.20	-19.00
4.70	3.90	3.10	2.30	1.50	0.70	-0.10	-0.90	-1.70	-2.50	-3.30	-4.10	-4.90	-5.70	-6.50	-7.30	-8.10	-8.90	-9.70	-10.50	-11.30	-12.10	-12.90	-13.70	-14.50	-15.30	-16.10	-16.90	-17.70	-18.50
5.20	4.40	3.60	2.80	2.00	1.20	0.40	-0.40	-1.20	-2.00	-2.80	-3.60	-4.40	-5.20	-6.00	-6.80	-7.60	-8.40	-9.20	-10.00	-10.80	-11.60	-12.40	-13.20	-14.00	-14.80	-15.60	-16.40	-17.20	-18.00
5.70	4.90	4.10	3.30	2.50	1.70	0.90	**0.10**	-0.70	-1.50	-2.30	-3.10	-3.90	-4.70	-5.50	-6.30	-7.10	-7.90	-8.70	-9.50	-10.30	-11.10	-11.90	-12.70	-13.50	-14.30	-15.10	-15.90	-16.70	-17.50
6.20	5.40	4.60	3.80	3.00	2.20	1.40	0.60	-0.20	-1.00	-1.80	-2.60	-3.40	-4.20	-5.00	-5.80	-6.60	-7.40	-8.20	-9.00	-9.80	-10.60	-11.40	-12.20	-13.00	-13.80	-14.60	-15.40	-16.20	-17.00
6.70	5.90	5.10	4.30	3.50	2.70	1.90	1.10	0.30	-0.50	-1.30	-2.10	-2.90	-3.70	-4.50	-5.30	-6.10	-6.90	-7.70	-8.50	-9.30	-10.10	-10.90	-11.70	-12.50	-13.30	-14.10	-14.90	-15.70	-16.50
7.20	6.40	5.60	4.80	4.00	3.20	2.40	1.60	0.80	**0.00**	-0.80	-1.60	-2.40	-3.20	-4.00	-4.80	-5.60	-6.40	-7.20	-8.00	-8.80	-9.60	-10.40	-11.20	-12.00	-12.80	-13.60	-14.40	-15.20	-16.00
7.70	6.90	6.10	5.30	4.50	3.70	2.90	2.10	1.30	0.50	-0.30	-1.10	-1.90	-2.70	-3.50	-4.30	-5.10	-5.90	-6.70	-7.50	-8.30	-9.10	-9.90	-10.70	-11.50	-12.30	-13.10	-13.90	-14.70	-15.50
8.20	7.40	6.60	5.80	5.00	4.20	3.40	2.60	1.80	1.00	0.20	-0.60	-1.40	-2.20	-3.00	-3.80	-4.60	-5.40	-6.20	-7.00	-7.80	-8.60	-9.40	-10.20	-11.00	-11.80	-12.60	-13.40	-14.20	-15.00
8.70	7.90	7.10	6.30	5.50	4.70	3.90	3.10	2.30	1.50	0.70	-0.10	-0.90	-1.70	-2.50	-3.30	-4.10	-4.90	-5.70	-6.50	-7.30	-8.10	-8.90	-9.70	-10.50	-11.30	-12.10	-12.90	-13.70	-14.50
9.20	8.40	7.60	6.80	6.00	5.20	4.40	3.60	2.80	2.00	1.20	0.40	-0.40	-1.20	-2.00	-2.80	-3.60	-4.40	-5.20	-6.00	-6.80	-7.60	-8.40	-9.20	-10.00	-10.80	-11.60	-12.40	-13.20	-14.00
9.70	8.90	8.10	7.30	6.50	5.70	4.90	4.10	3.30	2.50	1.70	0.90	**0.10**	-0.70	-1.50	-2.30	-3.10	-3.90	-4.70	-5.50	-6.30	-7.10	-7.90	-8.70	-9.50	-10.30	-11.10	-11.90	-12.70	-13.50
10.20	9.40	8.60	7.80	7.00	6.20	5.40	4.60	3.80	3.00	2.20	1.40	0.60	-0.20	-1.00	-1.80	-2.60	-3.40	-4.20	-5.00	-5.80	-6.60	-7.40	-8.20	-9.00	-9.80	-10.60	-11.40	-12.20	-13.00
10.70	9.90	9.10	8.30	7.50	6.70	5.90	5.10	4.30	3.50	2.70	1.90	1.10	0.30	-0.50	-1.30	-2.10	-2.90	-3.70	-4.50	-5.30	-6.10	-6.90	-7.70	-8.50	-9.30	-10.10	-10.90	-11.70	-12.50
11.20	10.40	9.60	8.80	8.00	7.20	6.40	5.60	4.80	4.00	3.20	2.40	1.60	0.80	**0.00**	-0.80	-1.60	-2.40	-3.20	-4.00	-4.80	-5.60	-6.40	-7.20	-8.00	-8.80	-9.60	-10.40	-11.20	-12.00
11.70	10.90	10.10	9.30	8.50	7.70	6.90	6.10	5.30	4.50	3.70	2.90	2.10	1.30	0.50	-0.30	-1.10	-1.90	-2.70	-3.50	-4.30	-5.10	-5.90	-6.70	-7.50	-8.30	-9.10	-9.90	-10.70	-11.50
12.20	11.40	10.60	9.80	9.00	8.20	7.40	6.60	5.80	5.00	4.20	3.40	2.60	1.80	1.00	**0.20**	-0.60	-1.40	-2.20	-3.00	-3.80	-4.60	-5.40	-6.20	-7.00	-7.80	-8.60	-9.40	-10.20	-11.00
12.70	11.90	11.10	10.30	9.50	8.70	7.90	7.10	6.30	5.50	4.70	3.90	3.10	2.30	1.50	0.70	-0.10	-0.90	-1.70	-2.50	-3.30	-4.10	-4.90	-5.70	-6.50	-7.30	-8.10	-8.90	-9.70	-10.50
13.20	12.40	11.60	10.80	10.00	9.20	8.40	7.60	6.80	6.00	5.20	4.40	3.60	2.80	2.00	1.20	0.40	-0.40	-1.20	-2.00	-2.80	-3.60	-4.40	-5.20	-6.00	-6.80	-7.60	-8.40	-9.20	-10.00
13.70	12.90	12.10	11.30	10.50	9.70	8.90	8.10	7.30	6.50	5.70	4.90	4.10	3.30	2.50	1.70	0.90	**0.10**	-0.70	-1.50	-2.30	-3.10	-3.90	-4.70	-5.50	-6.30	-7.10	-7.90	-8.70	-9.50
14.20	13.40	12.60	11.80	11.00	10.20	9.40	8.60	7.80	7.00	6.20	5.40	4.60	3.80	3.00	2.20	1.40	0.60	-0.20	-1.00	-1.80	-2.60	-3.40	-4.20	-5.00	-5.80	-6.60	-7.40	-8.20	-9.00

这个表格的第 i 行第 j 列表示：有 i 只黑羊，j 只白羊的时候，$ip - jq$ 的值（也就是 $(*)$ 式左边的值）。

如果第 i 行第 j 列是红色字体，就表示在有 i 只黑羊，j 只白羊的情况下，根据我猜测出来的结论，因为 $ip - jq < 0$ 了，所以管理者的最优策略是需要移走白羊的。如果字体是黑色则相反，不需要移走白羊。

如果第 i 行第 j 列是黄色背景的黑字，就表示根据上面提到的算法，管理者如果要采取最优策略，就必须要移走至少一只白羊。也就是说，黄色背景的位置，就是我的猜测和算法算出来的结果之间有出入的地方。如果认为算法是对的，并且不认为精度误差会出问题，那么我最开始的猜测就是错的。

进一步地，我观察了一下这个表中黄色的位置，有了另一个猜测，即实际上管理者的最佳策略应该是这样的：管理者选取了一个函数 $h(a, b)$，满足 $h(a, b)$ 略小于 $ap - bq$。一旦黑羊个数 a 和白羊个数 b 使得 $h(a, b) < 0$ 了，就不断移走白羊。

但是 $h(a, b)$ 具体是什么我就不知道了...... 不过，从结果上看，凭借 $ap - bq < 0$ 来作为一个策略应该已经算比较好了吧，尽管不是最优的 ~ 以我现有的能力只能解决到此，希望在以后的学习过程中能有新的发现。

运行动态规划算法的代码：https://paste.ubuntu.com/p/jtQmHrj8Cb/

运行模拟器的代码：https://paste.ubuntu.com/p/BNx9Qpfcv8/