Алексей Словеснов

Оптимальный алгоритм в игре быки-коровы.

Предисловие.

В этой статье рассматривается поиск оптимального алгоритма для игры быки-коровы. Будет доказано что,

1. Существует алгоритм позволяющий угадать любой загаданный номер, потратив на это не более семи ходов.
2. Не существует алгоритма позволяющего угадать любой загаданный номер, потратив на это не более шести ходов.
3. Найдено минимальное число номеров, угадываемых ровно за семь ходов, среди всех алгоритмов.

1 Детерминированные и недетерминированные алгоритмы.

Детерминированными алгоритмами назовём алгоритмы всегда делающие одинаковые ходы, зависящие только от ответов игрока. Например, первый ход в детерминированном алгоритме должен всегда быть один и тот же. Если же противник отвечает одна корова, то второй ход должен быть всегда одинаковым.

Недетерминированными алгоритмами назовём все остальные алгоритмы - они используют в своих ходах элемент случайности. Поэтому, например, первый ход, может быть любым.

Для всех дальнейших рассуждений нам будет достаточно рассматривать только детерминированные алгоритмы. Нетрудно понять, что для доказательства всех утверждений из предисловия достаточно рассматривать лишь детерминированные алгоритмы, так как если мы найдём оптимальный ход, то отклонение от него не улучшит алгоритм.

2 Алгоритм.

Во всех алгоритмах будем считать, что первых ход 0123. Это нисколько не ограничивает их общности. Сравнивая различные алгоритмы, мы будем добиваться, чтобы все номера угадывались не более чем за семь ходов, и число номеров, угаданных ровно за семь ходов, было бы как можно меньше. Так как общее число номеров 5040, то, для тестирования алгоритма, можно с помощью компьютерной программы проверить алгоритм по всем номерам, после чего получить статистику. В дальнейшем будет использоваться следующее обозначение ответов на ход, например, 2.1 будет означать, что получен ответ два быка и одна корова. Также будем называть семиходовкой номер, для угадывания которого, алгоритму требуется ровно семь ходов.

3 Сравнение алгоритмов.

В дальнейшем мы докажем, что нельзя создать алгоритм, при котором на угадывание любого номера можно потратить не более шести ходов, но есть алгоритмы которые угадывают любой номер не более чем за семь ходов. Поэтому можно считать, что любой алгоритм, у которого есть номера, которые он угадывает более чем за семь ходов, будет хуже любого алгоритма, который на угадывание любого номера тратит не более семи ходов. Но как сравнить два алгоритма, которые на угадывание любого номера тратят не более семи ходов. Будем считать, лучшим алгоритм, который угадывает наименьшее число номеров ровно за семь ходов.

4 Алгоритм максимального разбиения.

Алгоритм максимального разбиения состоит в следующем. Пусть уже сделано несколько ходов, рассмотрим группу номеров, которые удовлетворяют всем прошлым ходам. Теперь рассмотрим любой ход, после него группа элементов, удовлетворяющая прошлым ходам, будет разбита на несколько подгрупп. Два элемента попадают в одну подгруппу если они дают одинаковые ответы на последний ход. Среди этих групп найдём группу с максимальным числом номеров. Число номеров в максимальной подгруппе назовём числом данного разбиения. Теперь среди всех ходов найдём тот, у которого минимальное число разбиения - это и будет следующим ходом. Таким образом целью алгоритма максимального разбиения будет как можно сильнее раздробить группу номеров, удовлетворяющих всем предыдущим ходам.

Улучшения алгоритма максимального разбиения.

- Если удалось найти ход дробящий множество номеров так, что все группы содержат только один элемент, то
  1. Если этот ход удовлетворяет всем предыдущим ходам, то это самый лучший ход.
  2. Если этот ход не удовлетворяет одному из предыдущих ходов, то дальше нужно заняться поиском только среди номеров удовлетворяющих всем ходам.
- Если осталось только два номера, тогда лучшим ходом будет номер, имеющий меньший номер. Это ускоряет алгоритм, лучше хода уже не придумать.
- Если осталось только три номера, и на ход одного из них два оставшиеся дают разные ответы, то это лучший ход.
- Если есть два хода с одинаковым числом разбиения, тогда предпочтение будет отдаваться ходу, который удовлетворяет всем предыдущим ходам, так как он может быть загадан.

5 Алгоритм максимального дробления.

После нескольких ходов, следующий ход разбивает все оставшиеся номера на подгруппы. Каждая подгруппа будет содержать определённое число элементов $n_1, n_2, \dots$ Отсортируем эти числа по убыванию $n_1 \geq n_2 \geq \dots$ При разбиении группы двумя различными ходами, будет получено два набора номеров $n_1 \geq n_2 \geq \dots$ и $m_1 \geq m_2 \geq \dots$

Будем считать, что ход n дробит группу сильнее чем m, если $n_1 < m_1$ или $n_1 = m_1, n_2 < m_2$ и так далее. Например, набор $18, 18, 17 \dots$ лучше чем набор $18, 18, 18 \dots$ Алгоритм максимального дробления будет состоять в следующем - мы будем выбирать ход, который лучше всего дробит множество номеров, удовлетворяющих всем предыдущим ходам. Здесь верны те же улучшения, что и для алгоритма максимального разбиения, поэтому будем их использовать.

6 Подбор второго номера

После первого хода 0123, мы будем перебирать различные варианты второго хода, а затем использовать алгоритм максимального дробления. Выберем среди всех номеров второго хода тот, где число номеров, угаданных за семь ходов минимально. Перебор второго хода, значительно уменьшает число номеров, угадываемых ровно за семь ходов.

7 Сводные результаты.

После первого хода 0123 можно получить следующие ответы, в первой колонке указывается число номеров, во второй - ответ на ход

число номеров	ответ
1440	0.1
1260	0.2
720	1.1
480	1.0
360	0.0
264	0.3
216	1.2
180	2.0
72	2.1
24	3.0
9	0.4
8	1.3
6	2.2
1	4.0

После подбора второго номера удалось добиться того, что только при ответах 0.1 и 0.2 на первый ход остаются номера, которые угадываются ровно за семь ходов. Число номеров, угадываемых ровно за семь ходов, удалось значительно уменьшить.

Для сравнения, дана таблица с указанием числа семиходовок, для различных алгоритмов.

простой алгоритм максимального разбиения	648
улучшение 2 хода	374
приоритет оставшимся номерам	211
подбор второго хода+алг максим. дробления	76+23=99

В последнем алгоритме был достигнут результат, чтобы алгоритм решал 76 номеров ровно за семь ходов, если первый ответ даётся одна корова, и 23 номера, если первый ответ 2 коровы. Таким образом мы уже нашли алгоритм, который угадывает любой номер, затрачивая на это не более семи ходов.

8 Оптимальный алгоритм

Теперь займёмся поиском оптимального, алгоритма, который будет минимизировать число семиходовок. В дальнейшем нас будут интересовать только номера, дающие ответ 0.1 или 0.2 на первый ход, так как только среди этих групп номеров есть номера отгадываемые за семь ходов. Для первого ответа 0.1 компьютер нашёл, что вторым оптимальным ходом будет номер 2456, для 0.2 - лучший номер 4506. Вопрос выбора второго хода будет рассмотрен позднее. Второй ход опять будет разбивать группу на некоторые подгруппы. Выпишем в таблице подгруппы где число семиходовых номеров больше нуля, при этом мы считаем, что после второго хода мы пользуемся алгоритмом максимального дробления.

первый ответ 0.1	
ответ на 2й ход 2456	число семиходовок
0.1	55
0.2	17
1.1	2
1.0	1
0.0	1
всего	76

первый ответ 0.2	
ответ на 2й ход 4506	число семиходовок
0.1	17
0.2	5
1.0	1
всего	23

Дальнейшую оптимизацию необходимо производить только для указанных в таблице случаях.

9 Подбор третьего и четвёртого ходов.

Подбор третьего хода осуществляется следующим образом. Делается цикл по всем возможным третьих ходам, после этого для каждого возможного ответа делается алгоритм максимального дробления и отбирается ход минимизирующий число семиходовок. Кроме того, мы будем использовать верхнюю оценку на число семиходовок, если по некоторому номеру уже найдено, что число семиходовок больше либо равно чем оценка, то этот номер уже не лучший. Использование верхней оценки на число ходов позволит в дальнейшем сильно ускорить алгоритм перебора.

Программа нашла два вторых хода 2456 и 4506, для ответов 0.1 и 0.2, на самом деле эти номера эквивалентны. Мы берём одну из цифр первого хода, перемещаем её на новое место и добавляем три новых цифры, порядок и расположение которых, на самом деле, не имеют значения. В дальнейшем можем считать, что второй ход в обоих случаях 2456. Алгоритм максимального дробления даст на этом номере худший результат, если ответ на первый ход 0.1, но это связано с его несовершенством. В дальнейшем, при поиске оптимального алгоритма, мы уйдём от алгоритма максимального дробления. То, что вместо номера 4506 можно использовать номер 2456, будет доказано далее, в разделе о тождественных перестановках. А сейчас будем подбирать ходы 3 и 3-4, а после этого использовать алгоритм максимального дробления.

Для подбора хода 3 будем использовать следующее правило отбрасывания

1. Будем отбрасывать номера которые содержат цифру 0 на первом месте, либо цифру 1 на втором, либо 2 на третьем, либо 3 на четвёртом, то есть, чтобы коровы из первого хода не стояли на своих местах.

Для подбора ходов 3 и 4 будем использовать дополнительно второе правило отбрасывания, и множество третьих ходов.

2 Будем отбрасывать номера которые содержат цифру 2 на первом месте, либо цифру 4 на втором, либо 5 на третьем, либо 6 на четвёртом, то есть, чтобы коровы из второго хода не стояли на своих местах.

3 Третий ход будем делать только если это один из номеров: 7892, 7890, 7895, 7802, 7842, 7830, 7865, 7860, 7201, 7265, 7205, 7031, 7645, 7365, 7601, 1205, 5260, 1035, 4560, 1065.

первый ответ 0.1		
ответ на 2й ход 2456	подбор хода 3	подбор ходов 3-4
0.1	44	46
0.2	17	12
1.1	0	6
1.0	0	0
0.0	0	0

первый ответ 0.2		
ответ на 2й ход 2456	подбор хода 3	подбор ходов 3-4
0.1	17	12
0.2	2	0
1.0	0	0

Итак сейчас мы имеем только три случая, содержащих семиходовки. Пока что, лучшее найденное число номеров для ответа 0.1 равно 44+12=56, для 0.2 - 12, всего 68 номеров.

10 Подбор ходов 3-6

Правила подбора здесь используются как в правилах подбора ходов 3 и 4, но также будем проводить дополнительную оптимизацию по ходу 5. Шестой ход будет делать цикл по всем номерам. Если это ход из оставшегося множества и он разбивает начальное множество на группы по одному элементу, то сразу же можно сказать, что оценкой данного подмножества будет число номеров в подмножестве-1. Если ход не содержащийся в множестве разбивает подмножество на группы по одному элементу, то дальнейший поиск следует вести только среди номеров в группе. В противном мы не можем угадать все номера за 7 ходов. Сейчас алгоритм максимального дробления уже не используется. Результирующая таблица подбора ходов 3-6.

первый ответ 0.1		первый ответ 0.2	
ответ на 2й ход	семиходовки	ответ на 2й ход	семиходовки
0.1	43	0.1	12
0.2	12		

11 Убираем ограничения.

Что сейчас нам мешает сказать, что число семиходовок уже не удастся уменьшить? В процессе построения системы было использовано три правила, которые ограничивали общность.

1. Выбор второго хода 2456
2. Выбор третьего хода из множества: 7892, 7890, 7895, 7802, 7842, 7830, 7865, 7860, 7201, 7265, 7205, 7031, 7645, 7365, 7601, 1205, 5260, 1035, 4560, 1065.
3. Два правила отбрасывания номеров.

Сейчас мы будем постепенно отбрасывать наши ограничения и проводить расчёты. Проницательный читатель может спросить, а зачем было ставить ограничения, чтобы потом их отбрасывать? Дело в том, что, ставя себе некие рамки и проводя расчёты для них, мы, хотя и получали неоптимальные алгоритмы, тем не менее смогли сильно ускорять расчёты на компьютере и тем самым упростить себе дальнейшую жизнь. Например, проведя первые расчёты с весьма примитивными алгоритмами, мы получили, что нас интересуют только два варианта ответа на первый ход: одна корова и две коровы. Таким образом можно идти дальше отбросив все остальные ответы. В случае первых ответов 0.1 и 0.2 мы также получили верхние оценки для числа семиходовок, которую затем использовали в алгоритмах (теперь нас интересуют только номера с меньшим числом семиходовок). Итак, снимаем правила отбрасывания.

Результаты без правил отбрасывания.

первый ответ 0.1		первый ответ 0.2	
ответ на 2й ход	семиходовки	ответ на 2й ход	семиходовки
0.1	42	0.1	12
0.2	12		

12 Тождественные перестановки.

Чтобы двигаться дальше необходимо ввести новое понятие тождественной перестановки. Пусть у нас сделан один или несколько ходов, на которые уже получены ответы. Назовём тождественной перестановкой такое переобозначение цифр и перестановок позиций, что мы получаем тот же набор ходов с теми же ответами.

Рассмотрим, пример. Первый ход 0123 0.1, второй ход 2456 0.2. И рассмотрим перестановку цифр $1 \leftrightarrow 3$ $4 \leftrightarrow 6$ и замену мест $2 \leftrightarrow 4$. Нетрудно понять, что после такой замены номера 0123 и 2456 превращаются сами в себя. Иногда таких замен может быть больше, если первый ход 0123 0.1, второй 2456 0.1, теперь уже дополнительно номер 0123 может превратиться в номер 2456, при этом 2456 должен превратиться в 0123.

Перестановки неназванных цифр. В вышеприведённом примере, не были названы цифры 7, 8, 9. Совершенно очевидно, что если их переобозначить любым способом, то это будет тождественной перестановкой. Теперь ясно, что если два хода совпадают с точностью до тождественной перестановки, то из них достаточно рассмотреть лишь один. Использование тождественных перестановок позволяет отбросить очень много ходов.

В дальнейшем мы будем использовать следующее простое утверждение, попробуйте доказать его самостоятельно. Пусть есть первый ход t1 с ответом r1, и второй ход t2 с ответом r2 и есть преобразование t1 в t2 и одновременно t2 в t1, тогда число семиходовок для этих первых двух ходов, совпадает с числом семиходовок для первого хода t1 с ответом r2 и второго хода t2 с ответом r1. Пример, ходы 0123 0.1 и 2456 0.2 дают столько же семиходовок как и ходы 0123 0.2 и 2456 0.1.

13 Отбрасывание ограничения на выбор третьего хода.

Множество третьих ходов, которое мы брали, неполное, оно использовалось только для ускорения алгоритма и получения лучших оценок. В нём было отброшено большинство одиозных ходов, например 2017, но формально такие ходы могут дать лучшие результаты, поэтому для полной проверки множество необходимо расширить. Последний тест, исключает различные перестановки ещё не названных цифр 7,8,9 и использует тождественные перестановки. Результаты алгоритма.

первый ответ 0.1		первый ответ 0.2	
ответ на 2й ход	семиходовки	ответ на 2й ход	семиходовки
0.1	37 (4708)	0.1	10
0.2	10 (4061)		

Давайте посмотрим почему мы получили одинаковое число семиходовок для первого хода 0123 с ответом 0.1 и затем вторым ходом 2456 с ответом 0.2 и, наоборот 0123 с ответом 0.2 и затем 2456 с ответом 0.1. Дело в том что простой заменой цифр 3 на 6, 1 на 4, 5 на 0 с переменой мест 1 и 3, мы получим, что после обеих вариантов ходов, мы получим одинаковые множества (с точностью до вышеприведённых замен), а значит и число семиходовок у них одинаковое. Это связано с тем, что на оба первых ответа 0.1 и 0.2 мы делаем одинаковый ход. В дальнейших расчётах, пока мы будем считать, что второй ход мы делаем одинаковый, нам достаточно рассматривать только один из случаев.

Попробуйте найти лучший ход для 0123 0.2 и 2456 0.1 используя 0123 0.1 и 2456 0.2 и найденный лучший ход в этом случае 4061.

14 Множество вторых номеров.

На текущий момент был рассмотрен только второй ход 2456, пришло время перебрать все вторые ходы, для этого найдем все тождественные перестановки для первого хода 0123. Номера, с точностью до перестановок, для второго хода: 0124, 0132, 0134, 0145, 0214, 0231, 0234, 0245, 0456, 1032, 1034, 1045, 1204, 1230, 1234, 1245, 1435, 1456, 4567. Необходимо отметить, что для всех этих номеров найдены перестановки меняющие этот номер на 0123 и наоборот, в дальнейшем нам это пригодится. Всего таких номеров 19, один из них мы уже рассмотрели. Это 1456, который, с точностью до тождественных перестановок, совпадает с номером 2456. Кроме того у нас уже есть оценки сверху на первый ответ 0.1 47 ходов и на первый ответ 0.2 10 ходов. Каждый из этих номеров необходимо рассмотреть в качестве кандидата, на первый ответ 0.1 и на 0.2. В дальнейшем мы будем пользоваться таким соображением. Пусть мы рассматриваем некий номер, как кандидат на второй ход, после первого хода 0123 с ответом 0.1. После второго хода все номера, удовлетворяющие первому ходу, разобьются на несколько групп. Номера попадают в одну группу, если они дают одинаковый ответ на второй ход. Общее число семиходовок для выбранного второго хода будет равно сумме числа семиходовок, по каждой из подгрупп. Поэтому если мы найдём, что по одной или нескольким подгруппам сумма числа семиходовок больше или равна найденному нами выше числу 47, то этот ход будет заведомо не лучше чем 2456. Аналогичные рассуждения применимы к первому ходу 0123 с ответом 0.2, но уже с оценкой 10.

15 Перестановка цифр из первого номера.

Из группы вторых ходов выделим четыре номера, в которых присутствуют только цифры от 0 до 3. Это номера - 0132, 0231, 1032, 1230. Пусть был сделан первый ход 0123 и получен ответ 0.1, после второго хода могут быть получены только ответы 0.1 или 1.0. Если мы получим, что только для ответа 0.1 у нас будет 47 или больше семиходовок, то этот номер в качестве второго хода уже не годится. Тоже самое касается и первого хода 0123 с ответом 0.2, мы только посчитаем число семиходовок для второго ответа 0.2, если оно окажется 10 и больше, то это не самый лучший ход. Для всех этих номеров, как и ожидалось, улучшить оценку нельзя. Итак, осталось рассмотреть 14 номеров - 0124, 0134, 0145, 0214, 0234, 0245, 0456, 1034, 1045, 1204, 1234, 1245, 1435, 4567.

16 Первый взгляд.

Поскольку мы должны пробовать оставшиеся номера в кандидаты на первый ход 0123 0.1 и первый ход 0123 0.2, будет полезно первым делом рассмотреть 0123 0.1 и второй номер из группы с ответом 0.2, полученный результат можно будет применить также для ходов 0123 0.2 и второй номер 0.1, тем самым отбросив часть номеров как кандидатов, к тому же в этом случае мы сможем использовать более жёсткую оценку для оставшихся подгрупп. Вторым рассматриваемым вариантом будет 0123 0.1 и номер 0.1. Если сумма первых двух колонок будет ≥ 47, то номер не лучший на 0123 0.1, если же в первой колонке число ≥ 10, то номер не лучший на ход 0123 0.2. Результирующая таблица, для числа семиходовок, зачёркнутые номера уже не лучшие ходы для обоих случаев.

номер	0123 0.1 номер 0.2	0123 0.1 номер 0.1
~~0124~~	41	$\geq$6
~~0134~~	18	$\geq$47
~~0145~~	38	$\geq$47
~~0214~~	16	$\geq$32
~~0234~~	10	$\geq$47
~~0245~~	15	$\geq$47
~~0456~~	30	40
1034	4	$\geq$47
1045	9	38
1204	7	$\geq$47
1234	3	$\geq$47
1245	7	38
1435	7	$\geq$47
~~4567~~	$\geq$47	-

В результате мы исключили большинство номеров.

17 Кандидат на второй ход 0123 после первого ответа 0.2

Рассмотрим кандидата на второй ход 1234, после первого хода 0123 0.2, посмотрим какие могут быть ответы на второй ход 1234. Для облегчения жизни сначала запустим алгоритм с верхней оценкой 1 с оптимизацией третьего хода, (нас сейчас интересуют только ответы у которых потенциально могут быть семиходовки). После этого посчитаем для оставшихся групп, точное число семиходовок.

ответ	семиходовки, оптимизация 3го хода	семиходовки точное значение
0.1	$\geq$1	3 (ход 6305)
0.2	$\geq$1	1 (ход 3516)
1.0	$\geq$1	0 (ход 1056)
1.1	$\geq$1	0 (ход 1256)

Ход 1234 лучше, чем старый кандидат 2456, теперь число семиходовок, после хода 0123 0.2 удалось уменьшить с 10 до 4.

Теперь мы видим, что оставшиеся номера при новой оценке 4 уже не подходят как кандидаты на второй ход.

Примечание. Для номера 1034 с ответом 0.2, после первого хода 0123 0.2 точное значение семиходовок равно 1. То есть этот ход не походит даже как один из лучших.

18 Кандидат на второй ход 0123 после первого ответа 0.1

Теперь займёмся номером 1245 как кандидатом на второй ход, после первого хода 0123 0.1. Всё делаем также, как в предыдущем случае.

ответ	семиходовки, оптимизация 3го хода	семиходовки точное значение
0.1	$\geq$1	38 (ход 6078)
0.2	$\geq$1	7 (ход 5610)
0.0	$\geq$1	1 (ход 6730)
1.0	$\geq$1	0 (ход 6780)
1.1	$\geq$1	0 (ход 6543)

Итого, для второго хода 1245, после первого хода 0123 лучшее значение числа семиходовок 46. Новая, улучшенная оценка позволяет отбросить все остальные номера как кандидаты на второй ход.

19 Результирующее число семиходовок.

Итак, общее число семиходовок для первого хода 0123 0.1 - 46, для первого хода 0123 0.2 - 4. Лучшая оценка на число семиходовок равна 50.