



Борzych Алексей Каллевич

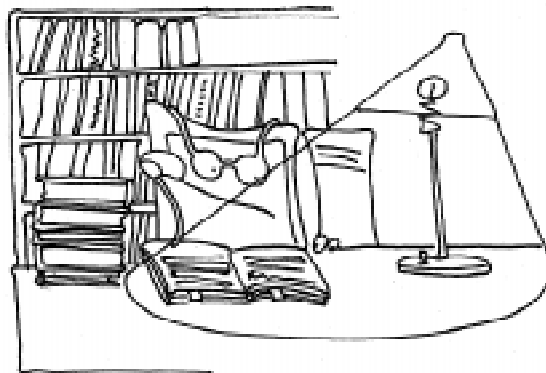
УНИВЕРСАЛЬНАЯ САМООБУЧАЮЩАЯСЯ МАШИНА ИЗ СПИЧЕЧНЫХ КОРОБКОВ

В этом номере мы публикуем работу одного из лауреатов конкурса «Петербургская Интернет-школа» 2000 года в номинации «Музей занимательной науки» Борzych А.К., создавшего виртуальный аналог самообучающейся машины из спичечных коробков. В последующих номерах мы познакомим читателей и с другими работами, ставшими победителями конкурса.

Может ли машина или, точнее, компьютерная программа самостоятельно обучаться в процессе какой-либо деятельности? Ответ на этот вопрос утвердительный. Простая самообучающаяся машина была описана несколько десятков лет тому назад и в популярной форме представлена в одной из статей книги Мартина Гарднера «Математические новеллы». В ней Мартин Гарднер предлагает в качестве материала для создания такой машины использовать спичечные коробки. Нам эта идея показалась настолько интересной, что мы решили создать виртуальный аналог этой машины, о котором и пойдет речь в этой статье.

Универсальная самообучающаяся машина из спичечных коробков работает очень просто. Пусть мы имеем некоторую пошаговую игру, в которой конечное множество возможных игровых позиций. В каждой игровой позиции существует конечное

множество возможных ходов, которые может сделать игрок. Для создания самообучающейся машины необходимо взять N коробков, где N – это количество возможных игровых позиций, и на каждом коробке нарисовать соответствующую игровую позицию. Также на каждом коробке надо показать множество возможных ходов для данной позиции и каждому возможному ходу присвоить определенный цвет (на разных коробках цвета могут повторяться). Далее следует внутрь каждого коробка положить по одной или несколько бусинок каждого цвета, обозначенного на коробке. После этого можно начинать игру человека с машиной.



Универсальная самообучающаяся машина из спичечных коробков...

Человек и машина ходят поочередно. Человек делает ход так, как считает более выгодным для победы. Машина делает ход следующим образом: она «берет» спичечный коробок, соответствующий данной игровой ситуации, хорошенько «встряхивает» его и случайным образом

«достаёт» одну из бусинок, кладя ее сверху на коробок. Какого цвета оказалась бусинка – такой ход и делает машина. Затем цикл повторяется снова.

Обучение происходит методом кнута и пряника. Если машина проигрывает, то ее «наказывают», удаляя те бусинки, которые лежат поверх коробков. Таким образом, уменьшается вероятность совершения в следующий раз последовательности ходов, приведших к проигрышу. Если машина выигрывает, то ее «поощряют», опуская обратно в коробки лежащие сверху бусинки и добавляя в каждый сыгравший коробок еще по одной (или несколько) бусинке того цвета, который «привел» к победе. Поэтому в новой партии игры вероятность совершения такой же последовательности ходов со стороны машины, что и в предыдущей, которая оказалась победной, немного увеличивается.

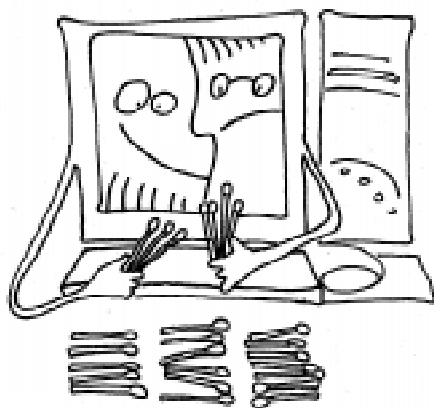
Следует отметить, что, если оба партнера играют идеально, не допуская «промахов» и «просчитывая» партию на достаточно большое количество ходов вперед, то исход любой игры всегда предопределен – обязательно победит либо тот, кто ходит первым, либо тот, кто ходит вторым (это зависит от игры). Поэтому если автомат имеет теоретическую возможность обыграть человека (преимущество первого хода на его стороне), то рано или поздно он научится это делать: чем больше партий сыграет человек с машиной, тем интеллектуальней будет играть машина, обучаясь на собственных ошибках и совершая их все реже и реже. После некоторого (возможно, большого) количества партий машина выработает идеальную стратегию игры, и человек окажется не в силах одержать победу! Понятно, что время обучения прямо пропор-

ционально зависит от комбинаторной сложности игры.

Интересно, что машина сможет обучиться игре даже и в том случае, когда преимущество первого хода не на ее стороне, но при обязательном условии, что человек иногда будет совершать «промахи» и делать не самые лучшие ходы – робот периодически будет их улавливать и улучшать свою стратегию игры. Однако, если человек никогда не будет допускать «промахов» и во всех партиях преимущество первого хода всегда будет на его стороне, то машина будет всегда проигрывать и никогда не сможет научиться безупречно играть!

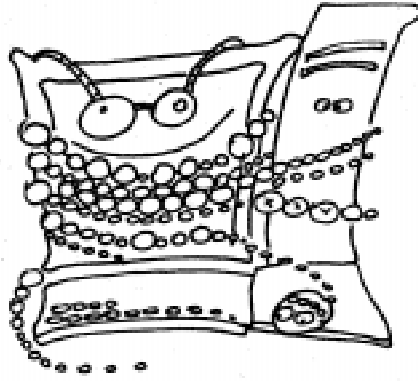
Особый интерес представляет реализация самообучающейся машины на компьютере и организация турниров между человеком и машиной, а между двумя компьютерами. Так как современные компьютеры обладают колоссальным быстродействием и имеют большие объемы памяти, то всего за одну секунду может быть сделано несколько тысяч или даже миллионов ходов! Компьютер способен достаточно быстро научиться идеально играть даже в очень сложные игры, просто немного поиграв «сам с собой».

Рассмотрим все вышесказанное на примере простой игры в «камушки». Правила игры таковы. Играют двое. На столе лежит некоторое количество предметов, например, камушков. За один ход можно взять один, два или три камушка. (В общем случае максимальное количество камушков, которое можно взять за один ход, может быть и другим.) Ходы делаются поочередно. Тот игрок, который берет последний камушек проигрывает партию. (В общем случае может быть и наоборот – правила игры могут варьироваться предварительно).



Компьютер способен ... научиться идеально играть даже в очень сложные игры, просто немного поиграв «сам с собой».

Если в начале игры на столе лежит N камушков, то для создания самообучающейся машины необходимо N спичечных коробков: каждому коробку соответствует одна возможная игровая позиция, которая определяется текущим количеством камушков на столе. Если за один ход можно брать



Исключительная ситуация...

от 1 до M камушков, то требуются цветные бусинки M различных цветов. Каждый цвет соответствует одному из возможных ходов: взять 1 камушек, взять 2 камушка, ..., взять M камушков. В коробок № 1, который отвечает за игровое положение, когда на столе лежит один камушек, следует положить некоторое количество (обозначим его X) бусинок цвета № 1, который отвечает за ход «взять один камушек». Бусинки других цветов в коробок № 1 не кладутся. В коробок № 2 следует положить X бусинок цвета № 1 и X бусинок цвета № 2, так как, когда на столе лежат два камушка, игрок может взять либо один из них, либо оба. В коробок № 3 – положить X бусинок цвета № 1, X бусинок цвета № 2 и X бусинок цвета № 3, ... В коробки № M , № $(M+1)$, ..., № N необходимо положить по X бусинок каждого цвета. После этого можно начинать игру.

Человек делает ходы из своих соображений. Машина делает ходы путем «встряхивания» коробков и «доставания» одной из бусинок. После окончания каждой партии происходит обучение машины: она либо «наказывается», либо «поощряется». При «наказании» в каждом «сыграншем» коробке уменьшается процентное содержание бусинок того цвета, который оказывается поверх коробка (по сравнению с бусинками других цветов). При «поощрении» – наоборот, увеличивается. Самый простой способ изменения этого процентного содержания – простое удаление «верхних» бусинок при «нака-

зании» или их «клонирование» (раздвоение) и обратное помещение в коробки при «поощрении».

Однако этот способ имеет один недостаток: может возникать как полное «опустошение» некоторых коробков (все бусинки кончаются), так и «переполнение» некоторых ко-

рбокков (число бусинок постоянно увеличивается). Для разрешения этой проблемы можно использовать стратегию обучения – «перекрашивание». Она заключается в следующем. Пусть в сыграншем коробке есть бусинки K различных цветов ($1 \leq K \leq M$). Цвет № L этого коробка «привел к победе» (цвет бусинки, лежащей поверх коробка). Тогда удалим из коробка по одной бусинке каждого цвета, кроме цвета № L , и добавим в коробок $(K-1)$ бусинок цвета № L , положив обратно также и верхнюю бусинку. В результате долевое содержание бусинок цвета № L увеличится, долевое содержание бусинок всех других цветов уменьшится, а суммарное количество всех бусинок останется неизменным. Аналогично, при наказании удаляем из коробка $(K-1)$ бусинок цвета № L и добавляем в коробок по одной бусинке всех остальных цветов. Исключительная ситуация: если число S бусинок цвета № L меньше $(K-1)$, то удалить $(K-1)$ бусинок цвета № L не получится, следовательно, удаляем все S бусинок цвета № L и добавляем S разноцветных бусинок каких-то произвольно выбранных цветов. Результатом применения стратегии обучения «перекрашивание» является постоянство количества бусинок в коробках. Также экспериментально установлено, что эта стратегия позволяет обеспечить гораздо более быструю «обучаемость» машины, чем стратегия «удаление-клонирование». Возможны и другие стратегии.

Идеальная тактика игры в «камушки» такова:

ПУСТЬ перед Вами лежит R камушков, причем Вы можете взять от 1 до T камушков.

ЕСЛИ по правилам последний взявший камушки проигрывает, то Вы должны взять $(R-1) \bmod (T+1)$ камушков. В некоторых случаях остаток от деления получается нулевым – это означает, что идеального хода не существует.

ЕСЛИ же побеждает последний взявший камушки, то для победы Вы должны взять $R \bmod (T+1)$ камушков ($A \bmod B$ – остаток от деления A на B).

Таким образом, если последний взявший камушки побеждает, то при любой прогрессивной стратегии обучения предельное заполнение бусинками коробков окажется таким: в коробках № 1, № $((M+1)+1)$, № $(2*(M+1)+1)$, ... будут только бусинки цвета № 1, в коробках № 2, № $((M+1)+2)$, № $(2*(M+1)+2)$, ... будут только бусинки цвета № 2, ..., в коробках № (M) , № $((M+1)+M)$, № $(2*(M+1)+M)$, ... будут только бусинки цвета № M , а вот в коробках № $(M+1)$, № $(2*(M+1))$,

№ $(3*(M+1))$, ... либо ничего не будет (если установлена стратегия «уничтожение-клонирование»), либо будут разноцветные бусинки (если установлена стратегия «перекрашивание»). Похожая картина наблюдается и для правил «последний взявший камушки проигрывает».

На сайте ЦПО «Информатизация Образования» по адресу www.ipo.spb.ru/internet-school/ch6.htm или на сайте автора по адресу www.borz.spb.ru/index2.html можно найти программное обеспечение, которое осуществляет полное компьютерное моделирование самообучающейся машины из спичечных коробков для игры в «камушки» (рисунок 1). Программный продукт обладает следующими возможностями:

1. Возможность задания правил игры. Вы можете указать начальное количество камушков, максимальное количество камушков, которое можно брать за один ход, задать установку «кто становится победителем» (последний взявший или наоборот), указать, кто первым ходит. Если вы не

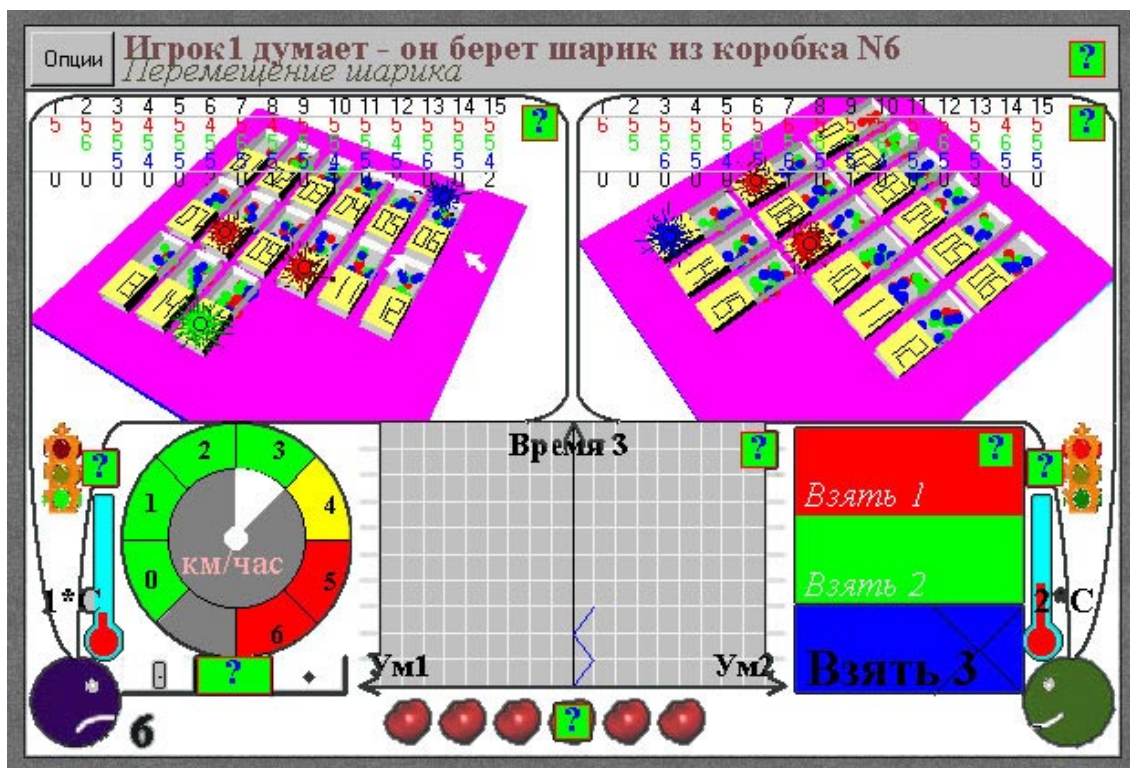


Рисунок 1

будете постоянно «ставить машину в тупик» заведомо проигрышными ситуациями, то какие бы правила Вы ни придумали, машина все равно к ним самостоятельно адаптируется, выработает идеальную тактику игры и будет Вас обыгрывать!

2. Возможность переключения в режим «компьютер против компьютера». В этом режиме друг с другом соревнуются два автомата. Чем больше они сыграют друг с другом, тем умнее будет каждый из них. У Вас есть возможность регулировать скорость их игры: на самой низкой скорости можно видеть медленное перемещение каждой бусинки, на самой высокой – за секунду они будут успевать сыграть друг с другом десятки партий! Также в любой момент Вы можете обратно переключиться в режим «человек против компьютера» и померяться силами с уже «натренированным» компьютером.

3. Возможность задания различных стратегий обучения и возможность задания параметра X (начальное количество бусинок каждого цвета в каждом коробке). Изменяя эти данные, можно влиять на скорость обучения. При одних настройках – машина «учится» очень быстро, при других – ей понадобится несколько сотен партий, чтобы «освоить» игру.

4. Наглядная трехмерная анимационная демонстрация всего происходящего с удобным пользовательским интерфейсом. Присутствует также график «зависимости ума компьютера от времени», который позволяет динамически наблюдать за успехами автомата.

Все, кого заинтересовала данная тема, могут обращаться к автору по E-mail: alex@borz.spb.ru или URL: borz.spb.ru.

НАШИ АВТОРЫ

*Борзых Алексей Каллевиц,
студент II курса СПбГЭТУ (ЛЭТИ)
кафедры «Математическое
обеспечение ЭВМ».*