

ОСОБЕННОСТИ РАЗРАБОТКИ ПОКЕРНОГО ИГРОВОГО ИНТЕЛЛЕКТА НА ОСНОВЕ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ.

Иванцов В.В., Зиганшин А.Т.
Томский политехнический университет
ivantsov.vladislav.v@gmail.com

Введение

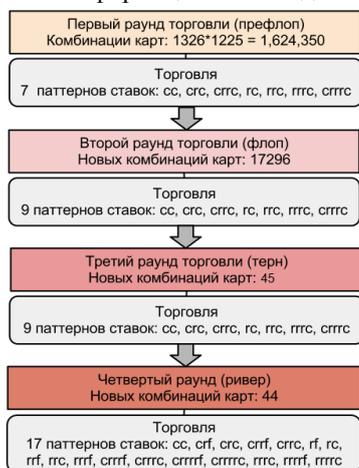
Задача разработки искусственного игрового интеллекта для покера является достаточно популярной, и, что самое главное, до сих пор нерешённой. Работы в данной области продолжаются уже не первый год, до сих пор не существует оптимального алгоритма, способного превзойти возможности человека, за исключением одного из частных случаев лимитированного покера (для двух игроков).

Сложность покера для машинного обучения

Техасский Холдем примечателен тем, что он является игрой с неполной информацией – каждый знает только свои карты и карты на столе (хотя все карты на столе открываются только на последнем круге прикупа). Также в нем присутствуют элементы случайности – карты игроков и на доске выметаются из колоды случайным образом. К этим двум аспектам игры в покер добавляется еще и третий – количество игровых состояний, которое в безлимитном Холдеме существенно выше, чем в лимитированном.

Рисунок 1 показывает структуру наиболее простого варианта игры, однако даже в простейшем случае игра содержит порядка 10^{18} игровых ситуаций, что делает данную игру практически не решаемой «в лоб», путём полного перебора всех возможных вариантов развития событий.

При создании покерных интеллектов обычно применяются различные упрощения, сокращающие количество игровых состояний. Но при этом возникает проблема потери существенной информации об исходной игре.



Схематическое представление игровых состояний Техасского Холдема. В обозначениях к паттернам ставок использованы следующие обозначения: 'c' - колл или чек; 'r' - бет, рейз, ревейз и т.д.; 'f' - флоп, сброс.

Рисунок 1. Структура игры с указанием новой информации

Концепция покерного интеллекта

Решение, рассматриваемое в данной статье, предполагает использование для обучения искусственной нейронной сети генетического алгоритма. Такой подход позволит опустить задачу разработки оптимальной покерной стратегии. Концепция генетических алгоритмов позволяет получить наиболее жизнеспособную особь путём селекционного отбора. В случае с покерным интеллектом жизнеспособность будет определяться эффективностью его игры. При этом интеллект не копирует поведение оппонента, а придерживается собственной стратегии, которая позволяет ему обыграть противника и получить наибольшую прибыль. Впрочем, сама стратегия полученного интеллекта будет поддаваться классификации, а скорее будет похожа на стиль игры человека, который был выработан благодаря полученному личному опыту.

Определение набора входных данных

Очень важным в процессе создания покерного интеллекта является получение от игровой среды наиболее полного и достаточного набора входной информации. Наличие избыточно информации усложняет процесс обучения ИНС, а нехватка информации может привести к принятию неверных решений.

Каждый игрок при игре в покер обладает следующей информацией: карты на руках, карты на столе, ставки игроков, количество денег на руках у каждого игрок, банк, а также расположение игроков за столом.

В первую очередь, следует убрать из набора входных данных информацию о картах. Вместо карт, комбинации которых достаточно сложно закодировать оптимальным образом для использования в ИНС, гораздо удобней и эффективней использовать вероятность выигрыша при известных картах на руках и на столе. Вероятность выигрыша может определена путём использования известных численных методов, например, метода Монте-Карло.

Ставки игроков и количество денег являются существенными данными. Исходя из этих данных, можно выработать оптимальную стратегию: например, можно попробовать задавить ставкой более слабого игрока, или же остережться сильного игрока в случае, если он начнёт повышать ставку. При этом сам банк на столе практически не играет никакой роли, поэтому им можно пренебречь.

Несущественной информацией является и расположение игроков за столом.

Структура искусственной нейронной сети

В связи со спецификой решаемой задачи, в данной работе используется каскад из двух классических нейронных сетей прямого распространения. Связано это с тем, что, по сути, основную задачу – принятие решения о ходе игрока, можно разделить на две подзадачи.

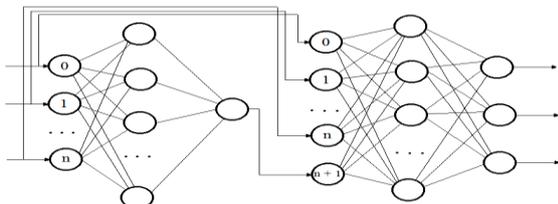


Рисунок 2. Каскад нейронных сетей

Первая подзадача – задача принятия решения о величине ставки, которая будет оптимальной исходя из текущего состояния игры. Такое решение принимается на основании силы комбинации, которая получается у игрока, и соотношения ставок и сумм доступных денег у игроков за столом. Опирируя этими данными, можно попробовать задавить игроков ставками, если сила руки получается небольшой, либо наоборот, попробовать поднять ставку так, чтобы не спугнуть противников.

Вторая подзадача сводится непосредственно к выбору хода игрока: сбросить карты, поддержать ставку либо повысить. При этом в качестве входных данных для анализа используется стандартный набор входных данных, а также результат работы предыдущей сети. Подобное разделение задач значительно упрощает процесс обучения ИНС.

Результаты

В исследовании были рассмотрены несколько интеллектов, отличающиеся структурами нейронной сети. Обучение осуществлялось с помощью генетического алгоритма, где в качестве индивида принимается полный набор весов каскада нейронных сетей. Эффективность обучения интеллекта определяется разнице между стартовым капиталом и капиталом после контрольной серии игр.

В ходе обучения интеллектов собирались промежуточные результаты, позволяющие построить график прогресса обучения. Данный график даёт возможность оценить качество обучения и выявить наиболее жизнеспособные интеллекты.

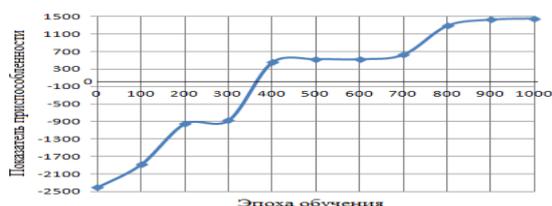


Рисунок 3. График процесса обучения интеллекта

График, представленный на рисунке 3, является одним из наиболее показательных для неэффективных решений. Как видно, прогресс интеллекта идёт неравномерно, сопровождается резкими скачками. Эти скачки вызваны не прогрессом в обучении интеллекта, а ошибками системы оценки эффективности обучения, когда интеллект выигрывает за счёт удачного стечения обстоятельств.

Подобные интеллекты характеризуются малым количеством скрытых слоёв в ИНС и малым количеством нейронов в них. Соответственно, подобные интеллекты неспособны обучиться таким сложным задачам, как игра в покер.

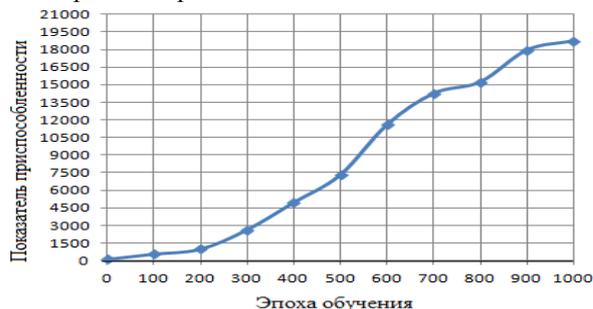


Рисунок 4. График процесса обучения интеллекта

Решения, подобные интеллекту на рисунке 4, являются наиболее эффективными. На графике видно, что показатель эффективности интеллекта достаточно монотонно возрастает, без резких скачков, что обусловлено параметрами генетического алгоритма с малыми силой и шансом мутации. В итоге решение получилось достаточно сильным, и к тому же имеет потенциал для дальнейшего обучения.

Подобные интеллекты характеризуются большим количеством скрытых слоёв и нейронов в них. Однако, следует отметить, что с возрастанием количества связей в ИНС, возрастает и сложность её обучения.

Заключение

Проведённое исследование позволило определить дальнейшие направления работы: необходимо модифицировать процесс оценки эффективности интеллекта. Также наибольший интерес представляют интеллекты, содержащие достаточно большое количество связей.

Литература

1. Искусственные нейронные сети. [Электронный ресурс]. – Режим доступа: <http://bigor.bmstu.ru/?cnt/?doc=NN/base.cou>, свободный.

2. Что такое искусственные нейронные сети? [Электронный ресурс]. – Режим доступа: <http://habrahabr.ru/post/134998/>, свободный.

3. Об искусственном интеллекте в покере.
[Электронный ресурс]. Режим доступа:
<https://geektimes.ru/post/173273/>, свободный

4. Генетический алгоритм. Просто о сложном.
[Электронный ресурс]. – Режим доступа:
<https://habrahabr.ru/post/128704/>, свободный.