

# СХОДИМОСТЬ ТЕКУЩЕЙ СТРАТЕГИИ АЛГОРИТМА $CFR+$

Взаимодействие двух игроков можно представить в виде игры с описанными правилами и результатами. В играх с нулевой суммой вариантом хорошей стратегии игры является равновесие Нэша. В данной работе исследуется сходимость текущей стратегии алгоритма  $CFR+$  в играх разного размера.

## ВВЕДЕНИЕ

Игры являются моделями проблем принятия решений, где чётко описаны правила, определяющие количество игроков, возможные действия, а также прибыль игроков в каждом возможном исходе игры. Игры, где в каждом исходе выигрыш одних игроков равен проигрышу других, называются играми с нулевой суммой.

В играх с неполной информацией игроки имеют знания о текущем состоянии игры, не доступные другим игрокам. Игрокам приходится принимать решения в условиях неопределённости в отличие от игр с полной информацией, таких как шахматы и шашки. Эта неполнота информации сильно затрудняет принятие хороших решений, так как игрокам необходимо думать о том, что думают другие игроки о состоянии игры, чтобы принимать качественные решения.

### I. ОПТИМАЛЬНЫЕ СТРАТЕГИИ И АЛГОРИТМ МИНИМИЗАЦИИ СОЖАЛЕНИЙ

В случае, если стратегии оппонентов полностью известны, можно построить контрстратегию, генерирующую максимальную прибыль против известных стратегий соперников. Однако такой подход практически не реализуем в реальной жизни. Первая проблема заключается в том, что стратегии оппонентов часто не известны. Вторая же проблема в том, что наилучшая контрстратегия перестает быть таковой, если оппоненты изменяют свою стратегию в процессе игры. Это может привести к тому, что выбранная стратегия станет худшей из возможных.

Равновесие Нэша – это набор стратегий, где стратегия каждого игрока является наилучшей контрстратегией против стратегий остальных игроков. Вместо точного расчета равновесных стратегий часто достаточно найти приближенное равновесие с точностью  $\epsilon$ , полученное с помощью итеративного алгоритма. Точность  $\epsilon$  определяется как разница между ожидаемой прибылью наилучшей контрстратегии и ожидаемой прибылью рассчитанной стратегии.

В 2007 году был разработан алгоритм *Counterfactual Regret Minimization (CFR)* [1] в ответ на большие затраты памяти предшествующих алгоритмов.  $CFR$  – итеративный алгоритм, играющий наедине с собой, который позволяет рассчитать  $\epsilon$ -равновесие Нэша в играх с нуле-

вой суммой для двух игроков. На каждой итерации алгоритм обходит игровое дерево со стороны каждого игрока. Решения в узлах принимаются в соответствии с «текущей» стратегией, которая обновляется на каждой итерации с помощью алгоритма *Regret Matching*. Результатом расчета  $\epsilon$ -равновесия является средняя стратегия, которая рассчитывается как среднее значение «текущих» стратегий на каждой итерации.

В результате развития был разработан алгоритм  $CFR+$  [2]. Данный алгоритм рассчитывает итоговую стратегию схожим образом, но теперь итерация  $t$  имеет вес  $t$  в результирующей стратегии. Также был изменен способ расчета текущей стратегии на каждой итерации с помощью алгоритма *Regret Matching+*. Эти изменения позволили рассчитывать стратегии как минимум на порядок быстрее.

### II. ИССЛЕДОВАНИЕ СХОДИМОСТИ ТЕКУЩЕЙ СТРАТЕГИИ АЛГОРИТМА $CFR+$

На практике текущая стратегия алгоритма  $CFR+$  часто сходится к оптимальной стратегии, однако теоретического доказательства сходимости нет. Для исследования сходимости текущей стратегии алгоритма  $CFR+$  были созданы 3 искусственные игры разного размера. Описание игр приведено в таблице 1.

Таблица 1 – Описание игр для исследования сходимости текущей стратегии алгоритма  $CFR+$

Название игры	Количество узлов принятия решений
Small Game	12
Average Game	6000
Big Game	1010000

На рисунках 1–3 представлена сходимость текущей и средней стратегий, рассчитанных алгоритмом  $CFR+$ . На графиках видно, что точность средней стратегии  $\epsilon$  увеличивается и сходится к 0 с увеличением количества итераций  $t$ . Точность текущей стратегии также увеличивается с увеличением количества итераций. Однако можно заметить, что сходимость текущей стратегии лучше в играх с большим количеством узлов. В играх «Average Game» и «Big Game» точность текущей стратегии до определенного момента даже выше, чем точность средней стратегии.

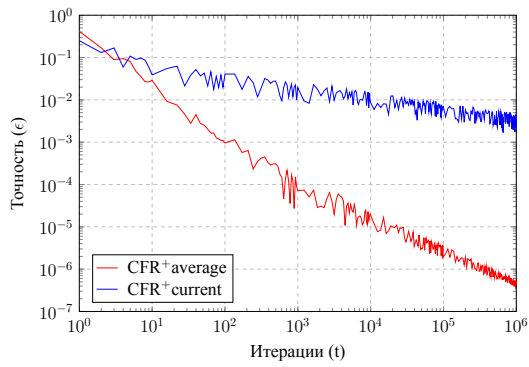


Рис. 1 – Сходимость стратегий в игре Small Game

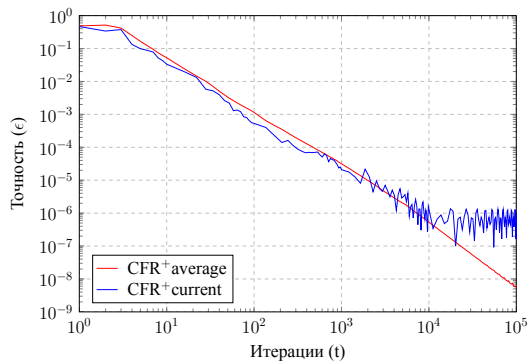


Рис. 2 – Сходимость стратегий в игре Average Game

### III. ВЫВОДЫ

Использование текущей стратегии в качестве результата позволяет не хранить среднюю стратегию во время расчета, что позволяет уменьшить требуемый объем оперативной памяти вдвое. Это в свою очередь позволяет рассчитывать оптимальные стратегии для игр большего размера.

*Пилипчук Валентин Олегович*, магистрант кафедры информационных технологий автоматизированных систем, valentinp.032@gmail.com

*Научный руководитель: Золоторевич Людмила Андреевна*, кандидат технических наук, доцент, zolotorevichla@bsuir.by

Количество итераций, необходимое для расчета оптимальных стратегий с необходимой точностью, может быть ниже, если использовать текущую стратегию в качестве результата, хоть и на исследованных играх эта разница была незначительна.

Стоит помнить, что сходимость текущей стратегии проверена на практике, но теоретических доказательств того, что она должна сходиться, не существует. Поэтому использование её в качестве результата возможно лишь в сценариях, когда возможно провалидировать полученный результат.

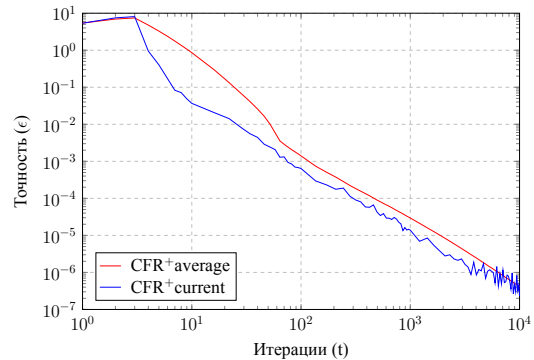


Рис. 3 – Сходимость стратегий в игре Big Game

### Список литературы

1. Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione. Regret minimization in games with incomplete information / Advances in Neural Information Processing Systems 20 (NIPS), pages 905–912, 2007.
2. Tammelin. O. Solving large imperfect information games using cfr+ / arXiv preprint arXiv:1407.5042, 2014.