

УДК 577.38*004.81

DOI 10.17726/phillT.2023.2.5



Эвристическое моделирование рефлексии в рефлексивных играх¹

Маркова Галия Муратовна,

*аспирант, ассистент, кафедра биофизики,
Институт фундаментальной биологии и биотехнологий,
Сибирский федеральный университет
Красноярск, Россия;*

*лаборант, лаборатория теоретической биофизики,
Институт биофизики Сибирского отделения
РАН – обособленное подразделение ФГБНУ Федеральный
исследовательский центр «Красноярский научный центр
Сибирского отделения РАН»
Красноярск, Россия*

GMarkova@ibp.ru

Барцев Сергей Игоревич,

*доктор физико-математических наук,
главный научный сотрудник, лаборатория теоретической
биофизики,*

*Институт биофизики Сибирского отделения
РАН – обособленное подразделение ФГБНУ Федеральный
исследовательский центр «Красноярский научный центр
Сибирского отделения РАН»
Красноярск, Россия;*

*профессор, кафедра биофизики,
Институт фундаментальной биологии и биотехнологий,
Сибирский федеральный университет
Красноярск, Россия*

bartsev@yandex.ru

¹ Работа поддержана грантом РФФ № 23-21-10041, Красноярского краевого фонда науки «Иерархия функциональных аттракторов в нейросетевых моделях рефлексии».

Аннотация. Функционирование субъекта в меняющейся среде наиболее эффективно с точки зрения выживания, если субъект способен формировать и сохранять внутренние представления о внешнем мире, а также использовать эти представления для принятия решений. Внутреннее представление (репрезентация) внешнего мира также называется рефлексией в широком смысле. Ее наличие позволяет побеждать в рефлексивных играх, т.к. внутреннее представление о противнике позволяет прогнозировать его поведение. В работе оценивается рефлексивный потенциал эвристических модельных объектов – искусственных нейронных сетей – в рефлексивных играх «Чет-нечет» и «Камень-ножницы-бумага». Использовались гомогенные полносвязные нейронные сети малых размеров (от 8 до 45 нейронов). Игры проводились между нейронными сетями с различными конфигурациями и параметрами (размер, величина шага модификации весовых коэффициентов). Представлен набор критериев рефлексивности, соотносящихся с разными уровнями рассмотрения: нейронным, поведенческим, формальным. С помощью нейронных сетей различных конфигураций показана транзитивность формального успеха в игре. Наиболее успешные конфигурации, тем не менее, могут не соответствовать другим критериям рефлексивности. Выдвигается гипотеза, что наилучшее соответствие критериям и, как следствие, универсальная успешность в задачах на рефлексии достижимы для гетерогенных конфигураций с иерархической структурой.

Ключевые слова: рефлексия; рефлексивная игра; эвристическое моделирование; искусственные нейронные сети; нейронная активность; аттракторы; стандартная эвристика.

Heuristic modeling of reflection in reflexive games

Markova Galiya Muratovna,

*Postgraduate Student, Assistant, Department of Biophysics,
School of Fundamental Biology and Biotechnology,
Siberian Federal University
Krasnoyarsk, Russia;*

*Laboratory Assistant, Laboratory of Theoretical Biophysics,
Biophysics Institute of the Siberian Branch of the RAS –
Division of Federal Research Center «Krasnoyarsk Scientific Center
of the Siberian Branch of the RAS»
Krasnoyarsk, Russia*

GMarkova@ibp.ru

Bartsev Sergey Igorevich,

*Doctor of Physical and Mathematical Sciences,
Chief Researcher, Laboratory of Theoretical Biophysics,
Biophysics Institute of the Siberian Branch of the RAS –
Division of Federal Research Center «Krasnoyarsk Scientific Center
of the Siberian Branch of the RAS»
Krasnoyarsk, Russia;*

*Professor,
Department of Biophysics,
School of Fundamental Biology and Biotechnology,
Siberian Federal University
Krasnoyarsk, Russia*

bartsev@yandex.ru

Abstract. The functioning of a subject in a changing environment is most effective from the point of view of survival if the subject can form, maintain and use internal representations of the external world for decision-making. These representations are also called reflection in a broad sense. Using it, one can win in reflexive games since an internal representation of the enemy allows predicting their future moves. The goal is to assess the reflexive potential of heuristic model objects – artificial neural networks – in the reflexive games “Even-Odd” (or “Matching pennies”) and “Rock-Paper-Scissors”. We used homogeneous fully connected neural networks of small sizes (from 8 to 45 neurons). Games were played between neural networks with different configurations and parameters (size, step size for modifying weight coefficients). A set of reflexivity criteria is presented, corresponding to different levels of consideration: neuronal, behavioral, formal. The transitivity of formal success in the game is shown. The most successful configurations, however, may not meet other criteria of reflexivity. We hypothesize that the best compliance with the criteria and, as a consequence, universal success in reflection tasks is achievable for heterogeneous configurations with a structure in which the formation of hierarchical systems of attractors is possible.

Keywords: reflection; reflexive game; heuristic modeling; artificial neural networks; neural activity; attractors; default heuristics.

Введение

В меняющейся среде с множеством объектов, каковой является окружающий нас мир, решение основных задач выживания (поиск пищи, партнеров, группы, защита от хищников и конкурен-

тов и т.д.) более эффективно, если субъект способен формировать в своем внутреннем пространстве представление о мире и его составляющих [1]. Сохранение таких представлений – память – позволяет сопоставлять новые сведения с уже имеющимися, принимать решения на этом основании и действовать. Способность формировать, хранить и использовать представления, согласно классификации, приведенной в работе [2], соотносится с уровнем сознания С1, что также говорит о тесной связи данной способности с феноменом сознания. Внутреннее представление (репрезентация) внешнего мира также именуется рефлексией в широком смысле. С научной точки зрения рефлексия представляет интерес, поскольку ее наличие или отсутствие можно зафиксировать в ходе определенных задач. Свойства, присущие субъектам, демонстрирующим рефлексию, и сопутствующие ее проявлениям, могут рассматриваться как играющие важную (и даже ключевую) роль в появлении сознания.

Одна из задач, в которой наличие рефлексии у испытуемого обеспечивает выигрышное поведение, – рефлексивная игра. Это взаимодействие между двумя или более участниками, преимущество в котором получает тот, кто точнее предскажет поведение противника и будет действовать, основываясь на этом предсказании. Рефлексия в такой задаче рассматривается как формирование внутреннего представления о противнике и учет этого представления в собственном поведении. Однако помимо рефлексивного представления в рефлексивной игре может применяться статистическая оценка встречаемости конкретных комбинаций ходов (регрессионная «модель» противника), стандартная эвристика и случайный выбор ходов.

Рассмотрим эти способы в отдельности.

1) накопление статистических данных о часто встречающихся комбинациях ходов (т.е. паттернах поведения) может быть реализовано, например, с помощью «гадалки» (машины) Шеннона [3]. Этот алгоритм подразумевает накопление сведений о количестве повторений определенных комбинаций ходов в ячейках памяти. «Решение» о текущем ходе машины принимается по наибольшей величине в соответствующих игровой ситуации ячейках. Мы считаем, что, хотя алгоритмы такого типа формируют своего рода «представление» о поведении противника, основанное на вероятности ходов, их функционирование не является рефлексивным.

Сведения о противнике должны накапливаться достаточно долго, они «инертны» и неуниверсальны, т. к. относятся к текущему поведению конкретного противника;

2) стандартная эвристика (default heuristics, «эвристика по умолчанию») – это распространенная стратегия, позволяющая упростить принятие решений. В случае рефлексивных игр стандартная эвристика проявляется как выбор хода по правилу «выиграл – повтори ход, проиграл – смени» [4]. Хотя с точки зрения биологии и здравого смысла такой подход разумен, он не является рефлексивным, т. к. поведение противника при выборе хода не имеет значения и представление о нем не формируется;

3) случайный выбор ходов также может оказаться эффективным на коротких игровых сериях, т. к. делает поведение игрока непредсказуемым для противника. Однако при длительном взаимодействии случайные ходы, в соответствии с равновесием Нэша, гарантируют не выигрыш, а ничью. Данный подход также не предполагает наличия представлений о противнике. Наконец, известно, что полностью случайный выбор ходов для игрока-человека невозможен.

Настоящая работа посвящена моделированию рефлексии как функции, проявляющейся в ходе рефлексивной игры. Мы используем эвристический метод, предложенный Дж. фон Нейманом [5] и основанный на использовании эвристических модельных объектов – формальных и сконструированных, соответствующих в определенном смысле широкому классу реальных объектов. Данный метод позволяет выявлять ключевые закономерности в функционировании систем независимо от их конкретной структуры и свойств. В качестве эвристических модельных объектов в работе применяются искусственные нейронные сети (ИНС). Еще одно методологическое основание нашей работы – биофизический подход к проблеме, суть которого состоит в рассмотрении наиболее простых объектов, задач и ситуаций для выявления наиболее общих характеристик рассматриваемого явления. В соответствии с этим, мы используем гомогенные ИНС малых размеров (от 8 до 45 нейронов) и рефлексивные игры с простейшими правилами взаимодействия.

Цель работы заключается в оценке рефлексивного потенциала эвристических модельных объектов – искусственных нейронных сетей – в рефлексивных играх «Чет-нечет» и «Камень-ножницы-

бумага». Формирование, хранение и использование представлений о противнике происходит в пространстве нейронной активности ИНС, которая определяет текущее состояние ИНС и их игровое поведение. Анализ нейронной активности позволяет выявить характеристики, сопутствующие рефлексивному поведению ИНС.

Материалы и методы

В работе в качестве основной конфигурации ИНС использовались рекуррентные (полносвязные) нейронные сети малых размеров (8, 15, 30, 45 нейронов), с двумя входами и двумя выходами (в игре ЧН) или тремя входами и тремя выходами (в игре КНБ), без отдельного входного и выходного слоя. Такая структура ИНС является простейшей, что удовлетворяет условиям нашего подхода. Вычисление сигналов на нейронах ИНС производится согласно формулам:

$$\alpha_i^{n+1} = \frac{\rho_i^n}{a + |\rho_i^n|}, \rho_i^n = \sum_j w_{ij} \alpha_j^n + A_i^n,$$

где w_{ij} – матрица весовых коэффициентов, A_i^n – входные сигналы, α_i^n – выходной сигнал j -го нейрона в n -ый момент времени, a – константа, задающая крутизну активационной функции нейрона. Начальные значения в матрице весовых коэффициентов задавались с помощью генератора случайных чисел в диапазоне (-0.025; 0.025).

Обучение ИНС производилось с помощью алгоритма обратного распространения ошибки. Для оценки качества функционирования ИНС использовалась квадратичная функция потерь:

$$C = \frac{1}{2} \sum_i (\alpha_i^n - \delta_i^n)^2,$$

где α_i^n – сигнал на выходных нейронах в момент времени n , δ_i^n – требуемый от сети сигнал в момент времени n .

Ранее было показано, что в рефлексивных играх рекуррентные (полносвязные) ИНС обладают преимуществом над слоистыми ИНС [6], на основании чего в дальнейших экспериментах применялись только рекуррентные конфигурации. Наличие обратных связей в структуре ИНС сопутствует наиболее эффективному (т.е. рефлексивному) игровому поведению, поскольку позволяет сохранять представления в пространстве нейронной активности.

Слоистые ИНС сохраняют сведения о предыдущих событиях в отдельном регистре – ячейках памяти, что сближает данные конфигурации с упомянутыми выше машинами Шеннона.

Для усиления обратных связей использовались также модифицированные конфигурации простых рекуррентных ИНС. В одном случае к имеющимся входам ИНС добавлялся еще один, на который поступал сигнал «1» или «-1» в соответствии с тем, выиграла данная ИНС или проиграла на предыдущем шаге игры. Предполагалось, что ИНС такой конфигурации будут способны оценить свою результативность, т.к. напрямую получают сигнал об успехе. В другом случае числовые сигналы, полученные ИНС на текущем шаге, обрабатывались не однократно, а дважды. Тем самым обеспечивалось наличие внутренних тактов функционирования ИНС, отличных от внешних – игровых. Такие ИНС способны выдать отклик именно на входной сигнал на текущем шаге, т.к. в течение внутренних тактов этот сигнал достигает выходов. ИНС без внутренних тактов нуждаются в двух игровых шагах, чтобы полноценно обработать полученный сигнал, в то время как отклик на него требуется от них мгновенно. Некоторые особенности функционирования модифицированных конфигураций ИНС описывались в работе [7].

Обозначения и свойства использованных модификаций ИНС приведены в таблице 1.

Таблица 1

Использованные в работе модификации ИНС

Название	Доп. вход	Временной режим
SRN (Simple Recurrent Network)	Нет	Такты функционирования совпадают с игровыми
SRN+	Да	
DTRN (Dual-Time Recurrent Network)	Нет	Есть внутренние такты
DTRN+	Да	

Рефлексивная игра «Чет-нечет» (ЧН) производится по следующим правилам. Игроки выбирают ход «0» или «1» независимо друг от друга, после чего их ходы сравниваются. Играющий за позицию «Чет» получает балл, если его ход совпал с ходом про-

тивника, играющий за позицию «Нечет» – если ходы не совпали. В случае ИНС, ход противника на предыдущем шаге игры становится входным стимулом на текущем шаге и подается на входы как «01» или «10».

Рефлексивная игра «Камень-ножницы-бумага» (КНБ) производится по следующим правилам. Игроки выбирают ход «0», «1» или «2» независимо друг от друга, после чего их ходы сравниваются. «0» (или «Камень») побеждает «1» («Ножницы»), «1» побеждает «2» («Бумага»), «2» побеждает «0». Аналогично ЧН, ход противника поступает на входы ИНС как «001», «010» или «100» на следующем шаге игры.

С точки зрения теории игр, ЧН является игрой с нулевой суммой, т.к. проигрыш одного игрока означает победу другого. Победитель получает балл на текущем ходе. В КНБ возможен еще один исход (ничья): когда игроки сделали одинаковый ход, тогда ни один из них не получает балл.

ИНС играли друг против друга в игры ЧН и КНБ. В режиме обучения все ИНС обладали модифицируемыми весовыми коэффициентами, начальные значения задавались с помощью генератора случайных чисел. В остальных режимах ИНС обладали фиксированными весовыми коэффициентами. Каждый набор характеристик играющих ИНС регистрировался для наборов из 500 шт. для статистической достоверности.

В работе [8] показано, что ИНС с конфигурацией типа SRN (т.е. наиболее простые из используемых) способны обучиться имитациям игр ЧН и КНБ. В этом случае противниками ИНС были четыре фиксированные последовательности ходов длины 12 тактов, имитирующие поведение квазипротивника. ИНС должны были действовать по игровым правилам, для чего требовалось быстро распознать, какая последовательность подается на вход в текущий момент, и сформировать подходящий отклик. Игровое взаимодействие ИНС с другими ИНС вместо фиксированных последовательностей – шаг к большему биологическому правдоподобию, т.к. оба игрока могут менять свое поведение в зависимости от ситуации, следовательно, потенциально способны к рефлексии в широком смысле.

Эксперименты проводились с помощью программных стендов, созданных нами для данного исследования в среде разработки Lazarus¹.

¹ <https://www.lazarus-ide.org>.

Результаты и обсуждение

При постановке экспериментов для достижения заданной цели мы столкнулись с рядом вопросов, требующих прояснения.

1. Какие условия будут в большей степени способствовать рефлексивному поведению: при допустимости или при запрете модификаций весовых коэффициентов (весов) ИНС?

Модификация весов происходит в режиме обучения и соответствует адаптивной подстройке самой структуры ИНС к текущей задаче. В таких условиях адаптивные возможности ИНС потенциально безграничны (т.к. веса могут становиться любыми, при достаточной продолжительности обучения), однако тогда осложняется сохранение представлений о противнике. Адаптивные возможности ИНС с фиксированными весами (после обучения) резко ограничены и определяются тем опытом взаимодействия, который был актуален в момент перед фиксацией, однако именно в таком режиме возможно отследить повторяющиеся паттерны поведения и нейронной активности ИНС, а также аттракторы – притягивающие множества в многомерном пространстве нейронной активности, к которым стремится состояние ИНС.

2. По каким критериям мы определим, что поведение играющих ИНС рефлексивно?

Поскольку рефлексия – наиболее эффективный способ решения задач типа рефлексивных игр, следует ожидать, что ИНС, демонстрирующие наличие рефлексии, будут успешнее прочих. Этот критерий был назван формальным и является основополагающим (какой смысл в рефлексии, если она не приводит к успеху?). Однако это не однозначный показатель рефлексивности, поскольку, как было сказано ранее, победа в игровой партии может достигаться и другими способами. Мы сформулировали набор критериев рефлексивности и на других уровнях рассмотрения помимо формального: поведенческом и нейрональном. Рассматривались следующие характеристики:

1) разнообразие игрового поведения. Учитывалось как количество комбинаций ходов длины 3 (ход ИНС – ход противника – следующий ход ИНС), встречающихся в игровой динамике ИНС хоть раз. Максимально возможное число комбинаций в игре ЧН – 8, в КНБ – 27. С точки зрения рефлексивности, предпочтительным являлось разнообразное поведение: если ИНС на протяжении всей

партии делает один и тот же ход («111...»), такое поведение нельзя назвать рефлексивным;

2) доля ходов, отклоняющихся от стандартной эвристики «выиграл – повтори ход, проиграл – смени». Учитывалось как количество использований комбинаций ходов, где после проигрыша ИНС следует смена хода, а после выигрыша ход повторяется. Предпочтительной являлась ситуация, когда в игровой партии более половины ходов отклоняются от стандартной эвристики;

3) вид аттрактора нейронной активности, возникающего при «свободном поведении» ИНС после одиночного иницирующего стимула. Данная характеристика определялась у ИНС с фиксированными весами вне игрового процесса и позволяла оценить стабильное внутреннее состояние данной ИНС. Возможные виды аттракторов: устойчивая точка, предельный цикл (произвольной длины), квазихаотический (неточечный непериодический аттрактор). Также определялся вид аттрактора, возникающего в ходе игры против другой ИНС, с фиксированными весами. Предпочтительными являлись квазихаотические аттракторы нейронной активности, поскольку известно, что функционирование на границе между упорядоченной и хаотической динамикой позволяет осуществлять вычисления оптимальным образом [9]. Кроме того, наличие квазихаотического аттрактора после иницирующего стимула свидетельствует о внутреннем состоянии ИНС, обладающем чувствительностью к малым изменениям среды. Такая система является менее «инертной», чем аналогичная с точечным или циклическим аттрактором;

4) корреляционная размерность аттракторов нейронной активности. Рассчитывалась по методу, описанному в [10], для ИНС во всех режимах функционирования: с модифицируемыми весами (обучение), с фиксированными весами, в «свободном поведении». Известно, что дробная размерность аттракторов свидетельствует о наличии хаотической компоненты в динамике системы. В свою очередь, точечные и циклические аттракторы имеют малые размерности, близкие к 0 или не определяемые данным методом. В связи с этим предпочтительными являются дробные размерности, превышающие 1.

Отмечено, что игровое поведение ИНС в двух режимах (с модифицируемыми и фиксированными весами) различается по всем критериям рефлексивности и описывается разными статистиче-

скими распределениями. На рисунке 1 показаны распределения баллов, полученных ИНС конфигурации SRN+ (15 нейронов, шаг модификации весов 0.01) в обеих играх и режимах.

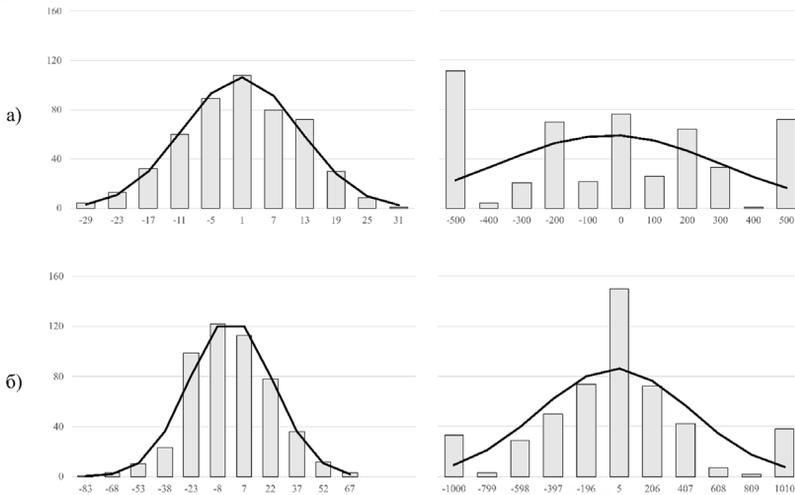


Рисунок 1. Распределения баллов, полученных наборами из 500 ИНС конфигурации SRN+ (15 нейронов, шаг модификации весов 0.01):

а) игра ЧН, слева направо: модифицируемые веса, фиксированные веса;

б) игра КНБ, аналогично. Гистограммы соответствуют экспериментальным данным, кривые – теоретическим (нормальное распределение)

Разнообразие возможностей возникает за счет начальной генерации весовых коэффициентов ИНС с помощью генератора случайных чисел, а в процессе игры система, которой является ИНС, обладает пороговой чувствительностью при формировании поведенческого отклика. Ход ИНС определяется исходя из того, на каком из выходных нейронов регистрируется больший сигнал, но не указывается минимально допустимая величина различия сигналов (т.е. нет варианта ответа «Пас»). Поэтому в системе, несмотря на детерминированность процесса после генерации начальных пара-

метров, возникают непредсказуемые (или, в терминах синергетики, катастрофические) события. Для описания подобных систем используется, например, модель песчаной кучи [11]. Известно, что для систем данного типа характерно наличие $1/f$ -шума, который также регистрируется в активности мозга, причем на разных уровнях организации [12]. Исходя из этого, у отдельных реализаций ИНС был проверен спектр нейронной активности в ходе рефлексивных игр (ЧН, КНБ) и получено, что $1/f$ -шум может появляться как у отдельных нейронов в составе ИНС, так и у совокупности всех нейронов (суммарная нейронная активность). Данный факт демонстрирует наличие еще одной «точки соприкосновения» между реальным объектом – биологической нейронной сетью, способной к рефлексии, и эвристическим модельным объектом – ИНС. Статистические показатели возникновения $1/f$ -шума у ИНС в ходе решения задач на рефлексии планируется определить в будущей работе.

3. Какова роль противника при формировании рефлексивного поведения?

Особенность рефлексивной игры как задачи для ИНС заключается в наличии второго «испытуемого» – ИНС-противника. Ходы противника напрямую учитываются при подсчете разнообразия поведения ИНС, т.к. включены в рассматриваемые комбинации ходов и влияют на их вариативность. Кроме того, существует представление о рангах рефлексии, согласно которому успешное рефлексивное поведение возможно только в том случае, если ранг рефлексии у игрока превышает таковой у противника (но не слишком сильно, т.к. иначе имеет место «overthinking» – чрезмерная обработка сигналов, не дающая выигрыша). Чтобы оценить вклад действий противника в поведение рассматриваемых ИНС, мы проводили игры между ИНС конфигураций SRN+ и DTRN+ (см. Материалы и методы) с фиксированным набором параметров (количество нейронов, шаг обучения) против ИНС тех же конфигураций, но разных размеров. Анализ по критериям рефлексивности, приведенным выше, показал, что почти все рассматриваемые характеристики ИНС (разнообразие поведения, доля отклоняющихся от стандартной эвристики ходов, размерность аттрактора в режиме: обучения против других ИНС, фиксации весов против других ИНС, после иницилирующего стимула после обучения против других ИНС) имеют близкие значения при

изменении размера ИНС-противников. Существенные изменения наблюдаются только в количестве квазихаотических аттракторов в режимах: против других ИНС с фиксированными весами, после иницирующего стимула после обучения против других ИНС (см. таблицу 2).

Таблица 2

**Количество квазихаотических аттракторов
(из 500 реализаций) при изменении размеров ИНС-игроков
и их противников**

	Кто с кем играл	SRN+		DTRN+	
		Против других сетей	«Свободный режим»	Против других сетей	«Свободный режим»
Игра «Камень-ножницы-бумага»	8vs8	0	16	0	14
	8vs15	139	22	34	8
	15vs15	9	284	0	278
	15vs30	5	284	0	268
	15vs8	3	262	0	252
	30vs15	115	350	20	340
	30vs30	194	356	19	346
Игра «Чет-нечет»	8vs8	1	10	0	123
	8vs15	80	8	22	96
	15vs15	1	19	0	175
	15vs30	2	24	0	159
	15vs8	2	33	0	200
	30vs15	17	91	3	329
	30vs30	10	68	1	291

В таблице 2 можно увидеть, что количество квазихаотических аттракторов при изменении размера противника меняется наиболее заметно для 8-нейронных SRN+ и DTRN+. Данная характеристика у 15- и 30-нейронных ИНС имеет достаточно близкие значения независимо от размера противников. Интерпретация этого феномена пока что неясна.

Также видно, что количество квазихаотических аттракторов во всех рассмотренных примерах больше у 30-нейронных ИНС, чем у 15-нейронных. Мы полагаем, что ИНС большего размера обладают большим потенциалом к разнообразному поведению в широком спектре игровых ситуаций, т.к. имеют большее внутреннее пространство для формирования и хранения представлений о поведении противника. Исходя из этого, увеличение количества квазихаотических аттракторов при увеличении размера ИНС в режиме игры против других сетей может говорить о том, что большие ИНС могут действовать неперiodично, несмотря на ограничения из-за фиксированных весов; а в «свободном режиме» после инципирующего стимула – свидетельствует о большей размерности внутреннего пространства ИНС.

4. Существуют ли универсальные характеристики, соответствующие рефлексивному поведению?

Подходя к данному вопросу, стоит рассмотреть также такую формулировку: рефлексивность – это свойство игрока (в данном случае, ИНС) или его поведения (которое определяется совокупностью характеристик игрового процесса)?

Если возможно выделить тип ИНС, универсально хорошо справляющихся с рефлексивными задачами, то логично полагать, что такие ИНС обладают собственными свойствами, обуславливающими этот успех. Рефлексивность их поведения в ходе решения задач является следствием этих свойств, и в таком случае мы можем говорить о рефлексивности (рефлексивном потенциале) самих ИНС.

Чтобы проверить возможность выделения таких ИНС, мы провели круговой турнир между ИНС конфигураций SRN, SRN+, DTRN, DTRN+ (см. Материалы и методы) разных размеров (15, 30, 45 нейронов), в режиме модифицируемых весов. Были подобраны оптимальные величины шагов модификаций весов для каждой конфигурации и размера, с помощью предварительных игр против референсной конфигурации (15-нейронные SRN с шагом модификации весов 0.003). Определено, что для большинства используемых конфигураций и размеров оптимален шаг 0.01.

Режим модифицируемых весов был выбран для кругового турнира, поскольку он обеспечивает, с одной стороны, максимальную способность ИНС к адаптации, а с другой – позволяет получить распределения характеристик, близкие к нормальному (см. рису-

нок 1). Турнирная таблица для игр ЧН и КНБ представлена в таблице 3.

Таблица 3

**Набранные баллы в результате кругового турнира
различных конфигураций и размеров ИНС,
в режиме модифицируемых весов**

Игра «Чет-нечет»			Игра «Камень-ножницы-бумага»		
Конфигурация и количество нейронов ИНС	Сумма баллов	Ошибка среднего	Конфигурация и количество нейронов ИНС	Сумма баллов	Ошибка среднего
DTRN+30	403	1	SRN+15	601	2
DTRN+45	401	2	DTRN+15	256	2
SRN+15	329	1	SRN15	227	2
DTRN+15	281	1	DTRN+45	139	2
DTRN45	67	1	DTRN30	123	2
SRN+45	-5	1	DTRN+30	79	2
SRN30	-50	1	SRN30	-102	3
SRN15	-64	1	SRN+45	-107	2
SRN+30	-134	1	DTRN15	-192	2
DTRN30	-186	1	SRN+30	-201	2
SRN45	-431	1	DTRN45	-344	2
DTRN15	-552	1	SRN45	-475	2

По турнирной таблице легко видеть, что наиболее успешными в круговом турнире в обеих играх оказались ИНС конфигурации SRN+ и DTRN+. Следовательно, мы можем предположить, что они обладают свойствами, способствующими возникновению рефлексивного поведения. Однако по результатам игр ИНС данных конфигураций против противников тех же конфигураций, но разных размеров, было определено, что характеристики ИНС на поведен-

ческом уровне рассмотрения (разнообразии поведения и доля отклоняющихся от стандартной эвристики ходов), а также размерности аттракторов нейронной активности на нейрональном уровне не во всех случаях соответствуют предпочтительным.

Мы полагаем, что при условии соответствия всем названным критериям рефлексивности игрок-ИНС может демонстрировать универсальный успех в играх с разными типами противников, в том числе, возможно, с другими алгоритмами и даже людьми. Однако приведенный выше результат свидетельствует о том, что по крайней мере среди рассмотренных конфигураций и параметров простейших гомогенных ИНС нет такой комбинации, которая позволяла бы получить универсально успешного игрока. Задаваясь вопросом, чего может не хватать для этого данному модельному объекту, мы предполагаем, что лучший результат можно получить при наличии способности формировать представления о противнике в разных временных масштабах. Такой объект должен обладать несколькими структурно обособленными «уровнями», в рамках каждого из которых строится свое представление. В этом случае станет возможным улавливать не только сиюминутные или, наоборот, наиболее масштабные тренды, но и некоторые промежуточные. Подобная иерархия представлений может анализироваться как система аттракторов нейронной активности разного уровня. Сложные системы аттракторов в настоящее время рассматриваются в биологических нейронных сетях [13].

Результаты кругового турнира также можно представить в виде схем, показывающих, над какими противниками каждая из конфигураций одержала победу (см. рисунок 2).

Схемы побед на рисунке 2 показывают наличие транзитивности у игроков-ИНС в данных рефлексивных играх (если игрок А побеждает игрока Б, а игрок Б побеждает игрока В, то игрок А побеждает игрока В). Такую цепочку можно отследить, например, в игре ЧН (рисунок 2а): SRN+15 побеждает SRN15, а SRN15 побеждает SRN45, при этом SRN+15 побеждает SRN45. Аналогичная цепочка присутствует и в игре КНБ (рисунок 2б). Этот факт также свидетельствует в пользу существования свойств, универсально способствующих проявлению рефлексии у обладающего ими объекта.

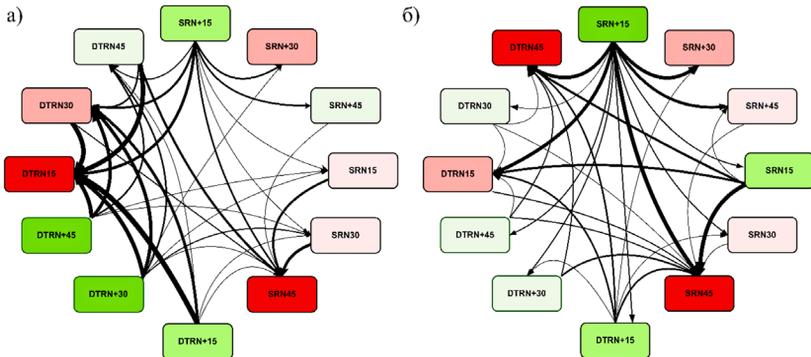


Рисунок 2. Схемы побед каждой из конфигураций, участвовавших в круговом турнире: а) в игре ЧН, б) в игре КНБ. Стрелка, ведущая от конфигурации А к конфигурации Б, показывает, что А побеждает Б. Толщина стрелки указывает на относительную величину выигрыша. Отсутствие стрелки между конфигурациями говорит о ничьей (суммарный балл близок к нулю). Цвет панелей с обозначениями конфигураций указывает на их место в турнирной таблице (темно-зеленый – первые места, темно-красный – последние, светлые оттенки – промежуточное положение)

Заключение

В настоящей работе, посвященной оценке рефлексивного потенциала эвристических модельных объектов – искусственных нейронных сетей (ИНС) – в рефлексивных играх «Чет-нечет» и «Камень-ножницы-бумага», приводится лишь часть наших наблюдений и рассуждений по данной теме. Показано, что рефлексивность игроков-ИНС необходимо оценивать по ряду критериев, относящихся к различным уровням рассмотрения: формальному, поведенческому, нейрональному. Среди конфигураций простейших ИНС малых размеров (от 8 до 45 нейронов) не удалось найти такую, которая обладала бы характеристиками, удовлетворяющими всем приведенным критериям рефлексивности. Требования по отдельным критериям при этом могут выполняться. В случае игры

ИНС против других ИНС, характеристики по большинству критериев практически не меняются. Исключение составляет формальный критерий (количество набранных баллов), а также количество ИНС в выборке, обладающих квазихаотическим аттрактором нейронной активности в режиме игры против других ИНС с фиксированными весовыми коэффициентами и в «свободном поведении» после одиночного иницирующего стимула. Наличие квазихаотического аттрактора внутреннего состояния объекта говорит о его чувствительности к малым изменениям стимулов, что представляется ценным для эффективного реагирования в условиях меняющейся среды. Подводя итог, мы заключаем, что рассмотренные нами эвристические модельные объекты в задачах, решение которых наиболее эффективно при наличии рефлексии, могут демонстрировать отдельные признаки рефлексивного поведения. Однако они, по-видимому, все же слишком просты для соответствия всем описанным критериям, что, по нашему предположению, являлось бы ключом к универсальному успеху в рефлексивных задачах. Дальнейшие исследования по данной теме мы планируем посвятить модельным объектам, способным к формированию иерархических систем представлений.

Литература

1. *Land M. F.* Do we have an internal model of the outside world? // *Philosophical Transactions of the Royal Society B: Biological Sciences.* – 2014. – Vol. 369, No 1636. – P. 20130045.
2. *Dehaene S., Lau H., Kouider S.* What is consciousness, and could machines have it? // *Robotics, AI, and Humanity: Science, Ethics, and Policy.* – 2021. – P. 43-56.
3. *Breazu M., Volovici D., Morariu D. I., Crețulescu R. G.* On Hagelbarger's and Shannon's matching pennies playing machines // *International Journal of Advanced Statistics and IT&C for Economics and Life Sciences.* – 2020. – Vol. 10, No 1. – P. 56-66.
4. *Wang Z., Xu B., Zhou H. J.* Social cycling and conditional responses in the Rock-Paper-Scissors game // *Scientific reports.* – 2014. – Vol. 4, No 1. – P. 5830.
5. *Фон Нейман Дж.* Теория самовоспроизводящихся автоматов. – М.: Мир, 1971. – С. 382. (Von Neumann J. *Theory of self-reproducing automata.* – М.: Mir, 1971. – P. 382.)
6. *Bartsev S., Markova G.* Recurrent and multi-layer neural networks playing “Even-Odd”: reflection against regression // *IOP Conference Series: Materials Science and Engineering.* – IOP Publishing, 2020. – Vol. 734, No 1. – P. 012109.

7. *Bartsev S.I., Markova G.M.* Does a Recurrent Neural Network Use Reflection During a Reflexive Game? // International Conference on Neuroinformatics. – Cham : Springer International Publishing, 2022. – P. 148-157.
8. *Markova G.M., Bartsev S.I.* Does a Recurrent Neural Network Form Recognizable Representations of a Fixed Event Series? // International Conference on Neuroinformatics. – Cham : Springer Nature Switzerland, 2023. – P. 206-213.
9. *Bertschinger N., Natschläger T.* Real-time computation at the edge of chaos in recurrent neural networks // Neural computation. – 2004. – Vol. 16, No 7. – P. 1413-1436.
10. *Grassberger P., Procaccia I.* Measuring the strangeness of strange attractors // Physica D: nonlinear phenomena. – 1983. – Vol. 9, No 1. – P. 189-208.
11. *Bak P., Tang C., Wiesenfeld K.* Self-organized criticality: An explanation of the $1/f$ noise // Physical Review Letters. – 1987. – Vol. 59. – P. 381-384.
12. *He B.J.* Scale-free brain activity: past, present, and future // Trends in cognitive sciences. – 2014. – Vol. 18, No 9. – P. 480-487.
13. *Zhang M., Sun Y., Saggat M.* Cross-attractor repertoire provides new perspective on structure-function relationship in the brain // Neuroimage. – 2022. – Vol. 259. – P. 119401.