**Федеральное государственное автономное образовательное учреждение высшего образования**

**КАЗАНСКИЙ (ПРИВОЛЖСКИЙ) ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ**

**ВЫСШАЯ ШКОЛА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И ИНФОРМАЦИОННЫХ СИСТЕМ**

Направление подготовки: 09.03.03 – Прикладная информатика

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**

**РАЗРАБОТКА СИСТЕМЫ ЭМОЦИОНАЛЬНОЙ ОЦЕНКИ НА ОСНОВЕ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ И НЕЙРОБИОЛОГИЧЕСКИ ИНСПИРИРОВАННЫХ МЕТОДОВ**

**Работа завершена:**

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_2016 г.

Студент группы 11-207 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Е.Ю. Майорова

**Работа допущена к защите:**

Научный руководитель

кандидат технических наук,

руководитель лаборатории машинного понимания

Высшей школы ИТИС

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_2016 г. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ М.О. Таланов

Директор Высшей школы ИТИС

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_2016г. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ А.Ф. Хасьянов

Казань – 2016 г.

ОГЛАВЛЕНИЕ

[ВВЕДЕНИЕ 4](#_Toc453863160)

[1. ЭМОЦИОНАЛЬНЫЙ ПРОЦЕСС ВНУТРИ ВЫЧИСЛИТЕЛЬНОГО АГЕНТА 6](#_Toc453863161)

[1.1 Роль эмоций у искусственного интеллекта 6](#_Toc453863162)

[1.2 Эмоциональные вычисления 7](#_Toc453863163)

[1.3 Эмоциональная оценка и ее связь с обучением 9](#_Toc453863164)

[2. ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕМ КАК СПОСОБ ЭМОЦИОНАЛЬНОЙ ОЦЕНКИ 11](#_Toc453863165)

[2.1 Принятие решений и обучение с подкреплением 11](#_Toc453863166)

[2.2 Общая постановка задачи обучения с подкреплением 12](#_Toc453863167)

[2.2.1 Значение оценочной обратной связи 12](#_Toc453863168)

[2.2.2 Постановка задачи 14](#_Toc453863169)

[2.3 Обзор методов решения задачи 16](#_Toc453863170)

[2.4 Метод временной разницы (TD-Метод) 17](#_Toc453863171)

[2.5 Выбор нейрональной архитектуры обучения с подкреплением 19](#_Toc453863172)

[3. РЕАЛИЗАЦИЯ СИСТЕМЫ ЭМОЦИОНАЛЬНОЙ ОЦЕНКИ НА ОСНОВЕ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ 21](#_Toc453863173)

[3.1 Нейросимулятор NEST 21](#_Toc453863174)

[3.2 Составление нейрональной модели 22](#_Toc453863175)

[3.3 Реализация модели 23](#_Toc453863176)

[3.4 Тестирование модели с различными данными 26](#_Toc453863177)

[3.5 Результаты 27](#_Toc453863178)

[ЗАКЛЮЧЕНИЕ 28](#_Toc453863179)

[СПИСОК ЛИТЕРАТУРЫ 29](#_Toc453863180)

[ПРИЛОЖЕНИЕ А 34](#_Toc453863181)

[ПРИЛОЖЕНИЕ Б 35](#_Toc453863182)

# ВВЕДЕНИЕ

В последние годы все большее количество людей привлекает такая область науки, как искусственный интеллект. Это утверждение доказывает множество примеров: AI2, CaptionBot, Jia-Jia, AlphaGo, Siri и т.д., - роботы, созданные с различными целями.

Существует гипотеза, что искусственный интеллект многое упускает, не реализуя эмоциональные процессы. [1] Роль эмоций у человека довольно велика. Например, они влияют на принятие решений и поведение. [2] Помимо этого, стоит отметить, что эмоции нужны для решения задач, включающих в себя социальное взаимодействие. [3] Воспроизведение эмоций необходимо для создания роботов, которые могут помогать человеку в его повседневной жизни. [3]

Ученые полагают, что теории возникновения и воспроизведения эмоций в естественном мозге могут оказать влияние на развитие роботов, обладающих эмоциями. [4]

Говоря об эмоциях, следует обратить внимание на процесс эмоционального оценки. Эмоции - филогенетический непрерывный механизм, гибко адаптирующийся к любым стимулам и ответам. [5] Гибкость эмоций получается путем отделения поведенческой реакции от события-стимула. Эмоциональный процесс - это оценка события по его значимости для индивида и предположение (прогнозирование) возможных действий, направленных на изменение отношений между человеком и окружающей средой. [6]

Основная проблема состоит в том, что современные вычислительные системы не могут проявлять эмоции, что в итоге приводит к их нежизнеспособности в условиях реального мира.

Объектом дипломной работы является эмоциональная оценка у искусственного интеллекта. В основе системы будет лежать процесс эмоциональной оценки у естественного интеллекта.

Актуальность темы подтверждается американским ученым Марвином Минским (1927 - 2016), одним из основателей Лаборатории искусственного интеллекта в Массачусетском технологическом университете. В своей книге "The emotion machine" он говорит о значимости эмоционального процесса для создания мыслящего искусственного интеллекта. [7]

В качестве системы для реализации эмоциональной оценки была выбрана система обучения с подкреплением. Это объясняется тем, что в настоящее время практически все методы прогнозирования у роботов полагаются на обучение с подкреплением. [8]

Цель дипломной работы заключается в разработке системы эмоциональной оценки на основе обучения с подкреплением. Эта система - один из этапов создания эмоционального искусственного интеллекта.

Для достижения данной цели необходимо решить ряд задач:

* изучить процесс эмоциональной оценки;
* изучить процесс обучения с подкреплением;
* проанализировать существующие описанные модели процесса обучения с подкреплением, выбрать наиболее подходящую под наши цели;
* составить нейрональную модель;
* изучить нейросимулятор "NEST";
* реализовать модель эмоциональной оценки с использованием нейросимулятора "NEST";
* протестировать полученную модель с различными параметрами и проанализировать полученные результаты.

# ЭМОЦИОНАЛЬНЫЙ ПРОЦЕСС ВНУТРИ ВЫЧИСЛИТЕЛЬНОГО АГЕНТА

## 1.1 Роль эмоций у искусственного интеллекта

Связь эмоций и познания имеет долгую историю в науке и философии. [9] В последние годы наблюдается растущий интерес к изучению эмоций, который объясняется, в частности, большим количеством новых знаний о мотивационных и эмоциональных процессов и их роли в познании. [4, 10]

Результаты исследования человеческой миндалины привели к пониманию взаимодействия эмоций и познания. Так, в 2006 году американским ученым Элизабет Фелпс был опубликован обзор, в котором рассматривалась связь между этими процессами. Фелпс выделила пять типов взаимодействия: эмоциональное обучение, эмоции и память, эмоции и восприятие, эмоции и социальные стимулы, изменение эмоциональное реакции. [11] Фелпс пришла к выводу, что научное понимание человеческого познания невозможно без эмоций. [11]

Так же выделяют следующие функции эмоций: оценка окружающей среды, система регулирования, подготовка действий, передача намерений, обдумывание и контроль. [12]

Эмоции являются неотъемлемой частью внутреннего состояния человека и оказывают влияние на совершаемый в различных ситуациях выбор. [13] В результате экспериментов, проведенных разными учеными в разное время [14, 15, 16, 17], мы можем утверждать, что эмоции, по всей видимости, влияют на вычисление значимости доступных действий, а так же, что это вычисление зависит от внутреннего состояния индивидуума. [13]

Многое можно сказать о роли эмоций в жизни человека. Но что, если мы имеем дело не с естественным, а с искусственным разумом? Нужны ли эмоции искусственному интеллекту?

Наличие у роботов эмоций может позволить стать социальному взаимодействию человека и робота более естественным. [18] Кроме того, эмоции необходимы для выживания конкретной особи и вида в целом. [19] Таким образом, роботу, разработанному для выживания в мире, потребуется эквивалентная живым организмам система. [19]

Как уже было сказано выше, эмоции – основная часть познавательных процессов. [20] К ним относятся воображение, внимание, восприятие, мышление, представление и др. [21] Так же когнитивистика изучает способность искусственного интеллекта к обучению в его окружающей среде. [22], следовательно, когнитивная архитектура развивается в направлении робототехники. Целью этого развития является познавательная способность искусственного интеллекта, схожая с аналогичной способностью человека. [22]

Таким образом, реализовав эмоциональный процесс у неживых организмов, мы можем получить некоторые преимущества перед роботами, которые не обладают эмоциями. Полностью реализованный эмоциональный процесс лежит за рамками данной работы. Сосредоточимся на эмоциональной оценке, которая является лишь частью эмоционального процесса в целом.

## 1.2 Эмоциональные вычисления

Для решения поставленной задачи необходимо понять, как можно реализовать эмоции у роботов.

Существует несколько подходов для создания искусственного интеллекта. Из них можно выделить две большие группы: искусственный интеллект по Джону Маккарти и биокомпьютинг (квазибиологическая парадигма). В первом случае утверждается, что искусственный интеллект не обязательно должен повторять структуру и протекающие в ней процессы, присущие естественному интеллекту. [23] В случае же биокопьютинга мы имеем в виду разработку искусственного интеллекта, структура, процессы или элементы которого протекают так же, как и у биологических организмов. Эти две группы не следует противопоставлять друг другу, так как Маккарти не утверждает, что воспрещается воспроизводить процессы, свойственные живым организмам. Исходя из сказанного в параграфе 1.1, мы склоняемся ко второй группе.

В частности, нас интересуют биологически инспирированные (или, другими словами, биологически вдохновленные) подходы. Нам близок подход, описанный американским ученым Марвином Минским, автором книги «The emotion machine», одним из основоположников эмоциональных вычислений. [7] Минский призывает понять работу нашего сознания, наши чувства и эмоции. Он говорит о создании искусственного интеллекта, близкого к естественному, сравнимого с человеческим. [24]

Общая модель, в которой объединяются наши взгляды на вычислительную систему и эмоции, опирается на модель куба эмоций Лёвхейма. Эта модель проводит связь между когнитивными процессами и нейробиологией. [25]

Лёвхейм отмечает связь между нейромедиаторами (дофамин, серотонин, норадреналин) и базовыми эмоциями, описанными ранее Томкинсом. К ним относятся страх, гнев, стыд, тоска, отвращение, удивление, радость, возбуждение. [26] Более подробная информация о кубе Лёвхейма изложена в приложении А.

Таким образом, реализация системы эмоциональной оценки будет проводиться с помощью биологически инспирированной нейронной сети.

## 1.3 Эмоциональная оценка и ее связь с обучением

Эмоции - филогенетический непрерывный механизм, гибко адаптирующийся к любым стимулам и ответам. [12] Эмоции являются адаптивной реакцией на какое-либо событие, значимость которого оценивается индивидом. [5] Понимание эмоциональной реакции влияет на будущие оценки. Теория эмоциональной оценки учитывает индивидуальные особенности, и, как следствие, эмоциональные реакции у различных индивидов на одно и то же событие. [27]

В данной теории существуют два основных подхода: структурный и процессный.

Структурный подход включает в себя изучение оценочного процесса и способов влияния различных оценок на испытываемые эмоции. Этот подход был подробно описан ученым Ричардом Лазарусом в 1991 году. [28]

Второй подход предполагает, что эмоциональная оценка – это непрерывный процесс. Этот подход учитывает, что эмоции имеют динамичный характер. В процессном подходе можно выделить две основные модели: двухпроцессная модель и многоуровневая модель Шерера.

К сожалению, существующие описанные модели являются высокоуровневыми, то есть не представлены в нейронном виде. С другой стороны, на основании процессного подхода можно сделать вывод, что эмоциональная оценка имеет предсказательный характер. Индивид решает, каким образом реагировать на событие в зависимости от множества параметров, таких как уместность последствий, способность с ними справиться и др. Таким образом, мы можем говорить о том, что индивид обучается эмоциональной оценке.

В настоящее время очень большое количество методов прогнозирования у искусственного интеллекта основано на обучении с подкреплением. [8] Обучение с подкреплением предполагает, что агент в процессе обучения получает награду (которая и является подкреплением), притом он стремится достичь максимизации этой награды. Более подробно обучение с подкреплением рассмотрено во второй главе.

# 2. ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕМ КАК СПОСОБ ЭМОЦИОНАЛЬНОЙ ОЦЕНКИ

## 2.1 Принятие решений и обучение с подкреплением

Поведенческая нейробиология решает различные вопросы, в числе которых вопрос принятия решений. Выбор оптимального действия основан на прогнозах долгосрочных последствий. [29] Например, 1) часто вкусу предшествует запах: почувствовав приятный аромат с кухни мы предполагаем, что вскоре сможем вкусно поесть; 2) мы не будем переходить дорогу даже на зеленый сигнал светофора, если на нас с высокой скоростью летит автомобиль.

Вопрос принятия решений был подробно исследован в поведенческой психологии на основе парадигмы Павлова. Так же появились и другие доказательства, что существуют области, контролирующие различные аспекты усвоенного поведения. В частности, множество доказательств свидетельствует о том, что нейромедиатор дофамин обеспечивает целевые структуры базальных ганглий фазовыми сигналами, которые передают ошибку прогнозирования, что в свою очередь влияет на обучение и выбор действия. [30 - 32] Так как вопрос принятия решений связан с прогнозированием, очевидно, что он связан и с обучением.

Современная интерпретация обучения с подкреплением исторически возникла из двух отдельных исследований. Первое исследование провели Ричард Саттон и Эндрю Барто, заинтересованные изучением создания искусственного интеллекта. Они разработали основные алгоритмы и концепции обучения с подкреплением. [33] Вторым исследованием занимались Димитрий Бертсекас и Джон Цициклис, специалисты в области исследования операций и оптимального управления. Их разработки заключаются в стохастических приближениях к методам динамического программирования, которые были названы нейродинамическим программированием. [34] Слияние этих двух направлений исследований сформулировало поведенчески вдохновленные алгоритмы эвристического обучения с подкреплением и предоставило инструменты для анализа свойств этих алгоритмов. [29]

Обучение с подкреплением - это обучение какого-либо агента действиям, чтобы максимизировать сигнала поощрения (вознаграждения), который принимает числовые значения. Обучение с подкреплением включает в себя следующие элементы: стратегия, функция вознаграждения, функция ценности и модель среды (в системах, имитирующих обучение с подкреплением). [35] Стратегия определяет поведение агента в каждый момент времени, функция поощрения является характеристикой вознаграждения, функция ценности характеризует состояние с учетом долгосрочных перспектив, а модель среды имитирует действия рассматриваемой среды. [35]

В процессе обучения агент взаимодействует со средой, получает от нее ответ, в зависимости от которого принимает дальнейшие действия. Каждое новое состояние характеризуется новым вознаграждением. Максимизация награды является целью обучения с подкреплением, для достижения которой могут использоваться различные алгоритмы: динамическое программирование, Прогнозирование и управление Монте-Карло, SARSA, TD-обучение и Q-обучение. [36 - 41] Все они были рассмотрены Саттоном и Барто. [35]

## 2.2 Общая постановка задачи обучения с подкреплением

### 2.2.1 Значение оценочной обратной связи

Как уже было сказано, цель обучения с подкреплением заключается в максимизации выгоды для интеллектуального агента. Здесь мы можем рассматривать две задачи: прогнозирование и управление.

В задаче прогнозирования обучение с подкреплением используется только для вычисления некой политики (стратегии), которая описывает для каждого посещаемого состояния ожидаемую в будущем награду. [8]

Задача управления ставит перед собой поиск набора стратегий, который максимизирует награду в режиме реального времени, т.е. во время переходов от одного состояния к другому. [8]

Для поиска оптимального маршрута используются алгоритмы, упомянутые в п. 2.1. При этом существует два вида нейрональных архитектур для прогнозирования и контроля: с обратной связью и без обратной связи. [8]

Классический пример обучения с подкреплением без обратной связи – собака Павлова. [42] В данном случае собака является интеллектуальным агентом, пища – безусловным раздражителем, звонок – условным. Условный раздражитель предсказывает безусловный, после обучения у агента появляется реакция на условный раздражитель. В данном случае нет обратной связи, т.к. реакция на условный раздражитель не влияет на представление стимулов. [8]

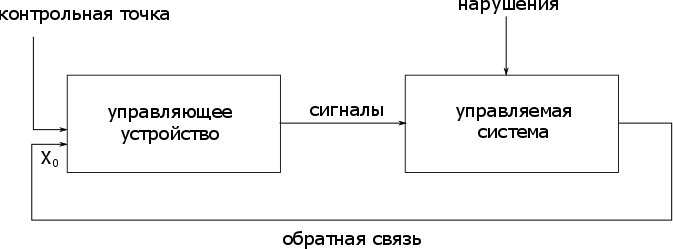
В случае существования обратной связи архитектура обучения представляет собой замкнутую схему. Замкнутость необходима для решения задачи управления, т.к. действия агента влияют на свои собственные входные значения. Эта схема изображена на рисунке 1. Точка X0 является сигналом обратной связи. [8, 35]

Рисунок 1 – Архитектура обучения с подкреплением с обратной связью

На рисунке 2 показана эта же архитектура, расширенная до так называемой модели Актёр-Критик (Actor-Critic). Эта модель получила такое название благодаря аналогии с актёром, который выбирает действие, и с критиком, который эти действия критикует. Критик представляет собой функцию оценочной стоимости. [8, 35]

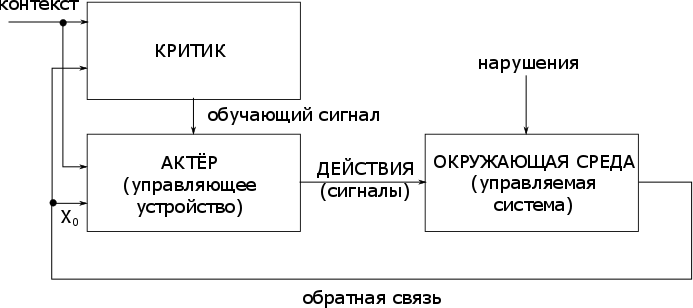


Рисунок 2 – Архитектура Актёр-Критик

Таким образом, обучение с подкреплением предполагает взаимодействие между некой средой и некой обучающейся системой, то есть наличие обратной связи.

### 2.2.2 Постановка задачи

Задачу обучения с подкреплением можно привести к математической модели. Рассмотрим, как она будет выглядеть в общем виде.

Пусть S – множество всех принимаемых состояний, и оно конечно.

В процессе обучения с подкреплением агент взаимодействует со средой. Другими словами, происходит игра между ними. Ниже приведен порядок действий в этой игре.

1. Инициализация стратегии и состояния среды
2. Для каждого момента времени
   1. Агент выбирает действие
   2. Среда генерирует награду
   3. Среда генерирует новое состояние
   4. Агент корректирует стратегию [43]

Эта игра будет называться марковским процессом принятия решений, если вероятность получения нового состояния и награды в нем зависит от текущего состояния и действия, совершенного агентом. [44]

Так как нам важна максимизация в долгосрочном периоде, то это требуется учесть.

Суммарная награда:

В общем случае суммарная награда является дисконтированной, т.е. формула (1) принимает вид:

где – коэффициент дисконтирования. Этот коэффициент был введен для того, чтобы обеспечить конечность суммы будущих наград, но он так же показывает, что люди и животные предпочитают ранние награды поздним. [29] Другими словами, чем больше тем агент будет более дальновидным.

Если – математическое ожидание при условии, что агент придерживается стратегии π, то [43]:

Функция ценности состояния при стратегии :

Функция ценности действия в состоянии при стратегии :

Таким образом, задача состоит в том, чтобы найти такую стратегию , при которой суммарная награда будет максимальна.

## 2.3 Обзор методов решения задачи

Саттон и Барто выделяют три фундаментальных метода решения задачи обучения с подкреплением: динамическое программирование, метод Монте-Карло и метод временной разницы (Temporal Difference, TD-метод). [35]

Методы динамического программирования предполагают, что модель окружающей среды идеальна и выполняется марковское свойство. Это предположение является следствием того, что применяются такие методы очень редко. Основная идея таких методов состоит в том, что функция ценности используется для создания и поиска оптимальных стратегий. [35] Алгоритм можно описать следующим образом:

1. Выбирается стратегия
2. До тех пор, пока стратегия не устойчива:
   1. Стратегия оценивается, – функция оценки стратегии
   2. Стратегия улучшается, – улучшенная стратегия,

Методы динамического программирования могут быть эффективными для определенных задач.

Методы Монте-Карло основываются на осредненной выгоде для некой выборки. Их используют для заданий, которые можно разбить на эпизоды. Здесь оценивание происходит не на каждом шаге, а после завершения очередного эпизода.

Методы Монте-Карло имеют преимущества перед методами динамического программирования. Например, они менее чувствительны к выполнению марковского свойства. [35]

Методы временной разницы в настоящее время являются наиболее важными среди всех методов обучения с подкреплением. [35] Они объединяют в себе идеи методов динамического программирования и Монте-Карло. Подробнее TD-метод рассмотрен в параграфе 2.4.

## 2.4 Метод временной разницы (TD-Метод)

Большинство алгоритмов, названных в параграфе 2.1, опираются на так называемый TD-метод – метод временной разницы. [8] Он был предложен Саттоном и Барто в 1990 году в качестве альтернативы модели Рескорла-Вагнера, которая учитывает торможение и блокирование, а также предполагает, что обучение происходит только тогда, когда награда не равна ожиданию. [45] У этой модели отмечаются некоторые недостатки. Например, она не может описывать события второго порядка: если стимул А предсказывает стимул B, а стимул B предсказывает состояние S, то очевидна связь между A и S, что никак не отражается в этой модели. [29] TD-метод устраняет недостатки модели Рекорла-Вагнера. [29]

Целью TD-метода является оценка ситуации с точки зрения будущих наград. Происходит это следующим образом: состояние и подкрепительный сигнал (награда) передаются из окружающей среды ко критику. Критик оценивает состояние функцией и высчитывает ошибку предсказания , которая используется для обучения и у критика, и у актёра. К актёру передается значение . Актёр на основании полученных данных генерирует стратегию . [29] На рисунке 3 изображена схема TD-обучения.

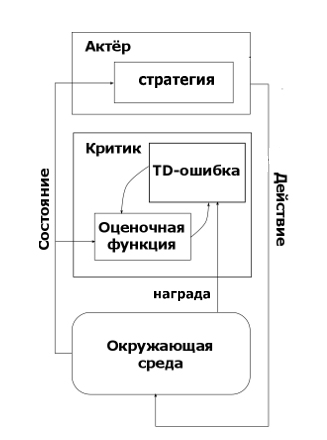
Рассмотрим этот процесс подробнее.

Рисунок 3 – TD-метод

Различные состояния из множества состояний следуют друг за другом согласно некоторому заранее определенному распределению *.* Награда наблюдается в состоянии с вероятностью *.* Значением состояния называется математическое ожидание суммы всех будущих наград.

Ожидание здесь является отношением между вероятностью перехода от одного состояния к другому и вероятностью вознаграждения в каждом состоянии. Тогда:

Согласованность между последовательными состояниями лежит в основе TD-метода. Последовательность имеет место лишь для тех значений, которые правильно прогнозируют ожидаемое . Если значения не верны, и равенство не выполняется, то говорят об ошибке предсказания временной разницы (TD-ошибка).

где – скорость обучаемости, а – TD-ошибка и

Таким образом, получаем:

С помощью TD-метода можно прогнозировать события, даже если среда не известна. [29]

## 2.5 Выбор нейрональной архитектуры обучения с подкреплением

До сих пор мы рассматривали лишь математические модели обучения с подкреплением. Однако, для нас представляет интерес то, как обучение с подкреплением происходит на нейрональном уровне.

Многие нейрональные архитектуры модели Актёр-Критик, описанные в литературе, сосредоточены на взаимодействии базальных ганглиев и коры. В этих архитектурах Критик описывается более подробно, чем Актёр. [8]

В настоящее время описаны следующие архитектуры:

* Параллельные взаимные архитектуры, модель Хоука
* Идеализированные параллельные взаимные архитектуры
* Расходящиеся взаимные архитектуры
* Параллельные невзаимные архитектуры
* Расходящиеся невзаимные архитектуры

Особое внимание стоит уделить схеме, предложенной Вёргёттеном и Порром, описывающей связь между TD-обучением и базальными ганглиями. [8] На рисунке 4 показаны эти связи с основными входами и выходами, а также отмечены возбуждающие и подавляющие связи между нейронами.

Рисунок 4 – Связь между TD-обучением и базальными ганглиями

Обозначения: ЧС(r) – черная субстанция (pars reticulata), БШ(in) – бледный шар (pars interna), БШ(ex) – бледный шар (pars externa), ЧС(c) – черная субстанция (pars compacta), ВОП – вентральная область покрышки, RRA – ретрорубальная область

Будем использовать данную схему для реализации системы.

В связи с трудностями перевода названий некоторых частей мозга на русский язык, все используемые нами описаны в приложении Б.

# 3. РЕАЛИЗАЦИЯ СИСТЕМЫ ЭМОЦИОНАЛЬНОЙ ОЦЕНКИ НА ОСНОВЕ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

## 3.1 Нейросимулятор NEST

Для моделирования процесса эмоциональной оценки будем использовать нейросимулятор NEST (The Neural Simulation Tool). Он был создан организацией NEST Initiaitive в 1994 году. Тогда он был выпущен под названием SYNOD. Позже его переименовали в NEST. Сейчас NEST является свободно распространяемым программным обеспечением, лицензия – GNU. [46]



Рисунок 5 – Логотип NEST Initiative

Фреймворк NEST был выбран для реализации модели по ряду причин. Во-первых, он поддерживает параллельные вычисления (PI, OpenMP), и его можно использовать на суперкомпьютерах. Во-вторых, у NEST есть пользовательский интерфейс PyNEST – пакет для языка Python. Это позволяет нейросимулятору быть простым в использовании. [47] В-третьих, особенности NEST, описанные ниже, подходят под наши цели.

Особенности NEST:

1. Система направлена на симулирование крупных структурированных сетей и биологически реалистичных элементов на разных уровнях описания;
2. Система использует итеративную инкрементную стратегию развития, поддерживающую работоспособность системы в любое время;
3. Система была разработана совместно несколькими группами заинтересованных разработчиков, она является исследовательским инструментом, и способна адаптироваться под требования исследователей. [48]

## 3.2 Составление нейрональной модели

Основная задача данной работы – построение высокоуровневой модели системы эмоциональной оценки на основе обучения с подкреплением. Нами был выбран метод временной разницы, а также нейрональная архитектура для его реализации.

Преобразуем схему так, чтобы ее использование в NEST не вызывало затруднений. Окончательный вариант представлен на рисунке 6.

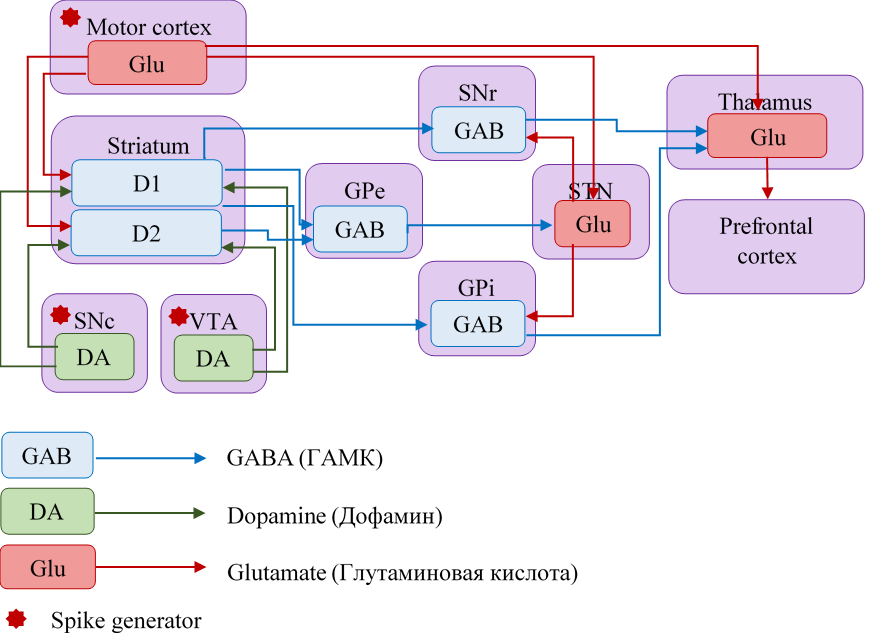


Рисунок 6 – Нейрональная архитектура

Спайковые генераторы (Spike generator) устанавливаются на двигательную кору, черную субстанцию (compacta), вентральную область покрышки.

Двигательная кора передает возбуждающие (Glutamate) сигналы в таламус, гипоталамус и полосатое тело. Под воздействием дофамина черная субстанция (compacta) и вентральная область покрышки передают сигнал полосатому телу. В свою очередь, полосатое тело подавляюще (GABA) воздействует на черную субстанцию (reticulata) и бледный шар. Бледный шар (externa) передает подавляющие сигналы гипоталамусу. Гипоталамус возбуждает черную субстанцию (reticulata) и бледный шар (interna), которые подавляюще воздействуют на таламус.

В итоге таламус получает сигналы от двигательной коры, черной субстанции (reticulata) и бледного шара (interna) и передает возбуждающий сигнал префронтальной коре.

## 3.3 Реализация модели

Для реализации нейрональной модели, описанной в параграфе 3.2 нам потребуются значения (см. параграф 2.4). Найдем пять первых значений с помощью формул (7 – 9). Результаты вычислений представлены в таблицах 1 – 3.

Таблица 1. Вычисление при

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| η | Шаг |  | | |
| 1 | 1 | 0,00 | -1,00 | -1,00 |
| 0,5 | 2 | -1,00 | -0,75 | -1,50 |
| 0,25 | 3 | -0,75 | -1,09 | -0,38 |
| 0,125 | 4 | -1,09 | -0,85 | -0,80 |
| 0,0625 | 5 | -0,85 | -1,11 | -0,33 |
|  |  |  |  |  |

Таблица 2. Вычисление при

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| η | Шаг |  | | |
| 1 | 1 | 0,00 | 0,00 | 0,00 |
| 0,5 | 2 | 0,00 | 0,00 | 0,00 |
| 0,25 | 3 | 0,00 | 0,00 | 0,00 |
| 0,125 | 4 | 0,00 | 0,00 | 0,00 |
| 0,0625 | 5 | 0,00 | 0,00 | 0,00 |
|  |  |  |  |  |

Таблица 3. Вычисление при

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| η | Шаг |  | | |
| 1 | 1 | 0,00 | 1,00 | 1,00 |
| 0,5 | 2 | 1,00 | 0,75 | 1,50 |
| 0,25 | 3 | 0,75 | 1,09 | 0,38 |
| 0,125 | 4 | 1,09 | 0,85 | 0,80 |
| 0,0625 | 5 | 0,85 | 1,11 | 0,33 |
|  |  |  |  |  |

Рассмотрим три случая, когда награда принимает разные значения: -1, 0 и 1. При , то есть при отсутствии награды (подкрепления), обучение не происходит. Если награда отрицательна, то есть по сути является наказанием, , обучение происходит, но тоже принимают отрицательные значения, следовательно максимизация выигрыша невозможна. При проведем эксперимент, используя соответствующие значения .

Представим вышеописанную модель на языке Python.

За связи на рисунке 6 отвечает следующий код:

1. connect(motor[Cortex], thalamus[thalamus\_Glu], syn\_type=Glu)
2. connect(motor[Cortex], striatum[D1], syn\_type=Glu)
3. connect(motor[Cortex], striatum[D2], syn\_type=Glu)
4. connect(motor[Cortex], stn[stn\_Glu], syn\_type=Glu)
5. connect(striatum[D1], snr[snr\_GABA])
6. connect(striatum[D1], gpi[gpi\_GABA])
7. connect(striatum[D1], gpe[gpe\_GABA])
8. connect(striatum[D2], gpe[gpe\_GABA])
9. connect(gpe[gpe\_GABA], stn[stn\_Glu])
10. connect(stn[stn\_Glu], snr[snr\_GABA], syn\_type=Glu)
11. connect(stn[stn\_Glu], gpi[gpi\_GABA], syn\_type=Glu)
12. connect(gpi[gpi\_GABA], thalamus[thalamus\_Glu])
13. connect(snr[snr\_GABA], thalamus[thalamus\_Glu])
14. connect(thalamus[thalamus\_Glu], motor[FrontalCortex], syn\_type=Glu)
15. if dopa\_flag:
    1. <…>
    2. connect(snc[snc\_DA], striatum[D1], syn\_type=DA\_ex)
    3. connect(snc[snc\_DA], striatum[D2], syn\_type=DA\_in)
    4. connect(vta[vta\_DA0], striatum[D1], syn\_type=DA\_ex)
    5. connect(vta[vta\_DA0], striatum[D2], syn\_type=DA\_in)

Выше был описан порядок воздействия разных структур мозга друг на друга. Здесь мы видим, что действительно, сначала воздействует двигательная кора (1 – 4), затем полосатое тело воздействует на черную субстанцию и бледный шар (5 – 8), бледный шар (externa) – на гипоталамус (9). Гипоталамус воздействует черную субстанцию (reticulata) и бледный шар (interna) (10 – 11), которые воздействуют на таламус (12 – 13). В конце таламус воздействует на префронтальную кору (14). Дофаминовое воздействие вынесено в отдельный блок (15): черная субстанция (compacta) и вентральная область покрышки воздействуют на полосатое тело (b – e). Типы связей (возбуждающая, подавляющая, воздействие дофамина) так же указаны, за них отвечает параметр syn\_type.

Значения используются при запуске генератора. Он генерирует спайки по пуассоновскому распределению. Действует с определенной силой, измеряемой в mV, т.е. по сути, генератор заряжает нейрон и повышает мембранный потенциал. Так же у генератора есть частота. На языке Python для него используется соответствующее условие generator\_flag. Рассмотрим код подробнее:

1. if generator\_flag:
2. connect\_generator(motor[Cortex], 0.1, T, rate=300, coef\_part=1, weight=2)
3. delta = [1.0, 1.5, 0.38, 0.8, 0.33]
4. k = 9.9
5. iter = 0
6. for i in range(5) :
   1. connect\_generator(snc[snc\_DA], k, k + 3.3, rate = 300, weight=delta[iter], coef\_part=1)
   2. connect\_generator(vta[vta\_DA0], k, k + 3.3, rate = 300, weight=delta[iter], coef\_part=1)
   3. iter += 1
   4. k += 3.3

Подключаем генератор к двигательной коре (2), время действия – от 0,1 мс до T (задается, в нашем случае T = 40 мс), частота – 300, параметр coef\_part отвечает за процент нейронов, которые будут подцеплены детектором (используется для отображения результатов в виде графиков), сила генератора weigh равна 2. Затем нужно подключить еще два генератора, но их силы зависят от , объявляем их (3). Чтобы дать генераторам время на «разгон», время действия поставим от 9,9 мс (4). Далее запускаем генераторы в цикле, будем менять силу, зависящую от , на каждом шаге. Один шаг длится 3,3 мс (особенность системы).

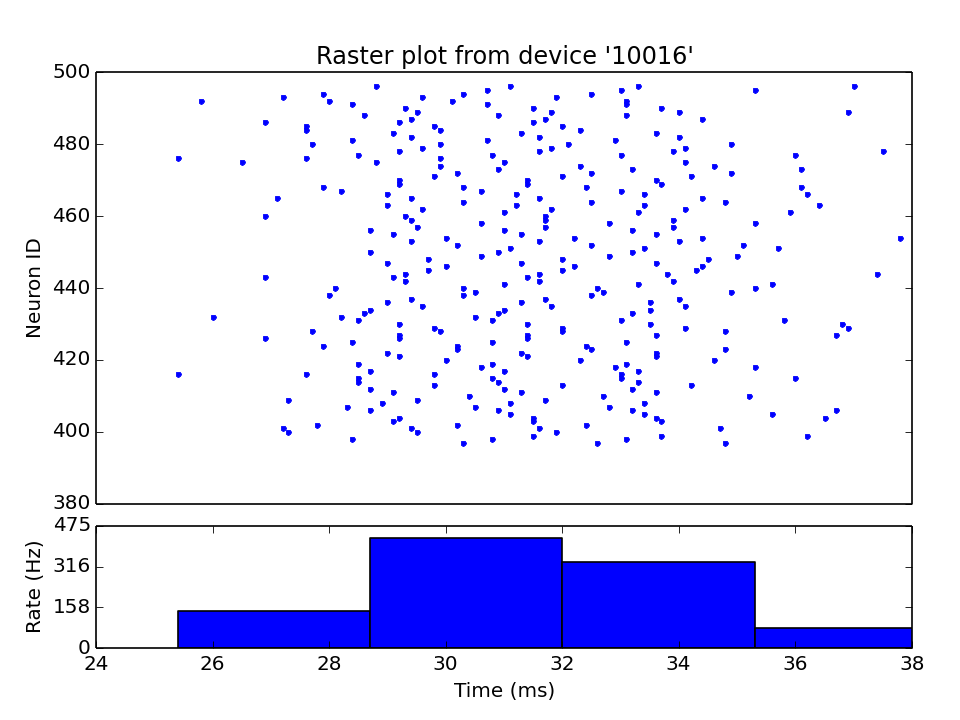
## 3.4 Тестирование модели с различными данными

До сих пор мы говорили лишь о воздействиях различных структур мозга друг на друга. Однако, зачастую наличия воздействия недостаточно. Например, воздействие может быть слабым и совершенно незначительным. Но у этой проблемы есть решение: можно определить синаптическую силу этого воздействия для любой связи. Такая сила называется весом, обозначается в программе с помощью параметра weight\_coef.

Нам нужно настроить веса таким образом, чтобы была видна активность структур мозга, участвующих в обучении с подкреплением.

На рисунке 7 показаны спайки на таламусе до настройки системы.

Рисунок 7 – Спайки на таламусе до настройки системы



## 3.5 Результаты

# ЗАКЛЮЧЕНИЕ

О создании эмоционального искусственного интеллекта ученые начали говорить еще в середине XX века, но направление «Эмоциональные вычисления» оформилось лишь в конце 90-х годов, когда профессор Массачусетского технологического института Розалинд Пикард выпустила книгу с одноименным названием. [49]

Развитие науки привело к попыткам создания роботов, обладающих эмоциями. Существуют разные подходы к этой задаче. Научно-исследовательская лаборатория «Машинное понимание» Высшей школы ИТИС Казанского Федерального Университета поставила перед собой цель построить машинное сознание, опираясь на нейробиологически инспирированные подходы. Была предложена когнитивная архитектура NEUCOGAR и поставлена задача реализовать ее.

Цель данной работы – создание системы эмоциональной оценки на основе обучения с подкреплением – этап в разработке эмоционального искусственного интеллекта. Эмоциональная оценка неразрывно связана с прогнозированием, которое, в свою очередь, связано с обучением.

В результате анализа литературы и различных исследований была составлена и реализована нейрональная модель. В результате симуляции данной модели были получены графики, показывающие активность структур мозга, участвующих в процессе их воздействия друг на друга. Так как воздействие не может быть однозначным, а каждая связь обладает синаптической силой, то такую силу, называемую весом, нужно настроить.

…

Создание искусственного интеллекта на основе биологически инспирированных подходах приведет к возможности создания эмоциональных машин. Эмоциональная основа позволит роботам познавать, накапливать опыт и выживать в условиях реального мира.

# СПИСОК ЛИТЕРАТУРЫ

1. Эмоциональные вычисления [Электронный источник] / Таланов Максим // Режим доступа: <http://postnauka.ru/video/45297> – свободный
2. The role of affect in decision making [Текст] / Lowenstein, G., Lerner, J.S. // In R. Davidson, K. Scherer, & H. Goldsmith (Eds.), Handbook of affective science, pp. 619-642. New York: Oxford University Press, 2003
3. Эмоциональный искусственный интеллект [Электронный источник] / Таланов Максим // Режим доступа: http://postnauka.ru/video/45296 – свободный
4. On the Role of Emotion in Embodied Cognitive Architectures: From Organisms to Robots [Текст] / Tom Ziemke, Robert Lowe // Springer Science+Business Media, LLC 2009
5. A systems approach to appraisal mechanisms in emotion [Текст] / David Sander, Didier Grandjean, Klaus R. Scherer // Geneva Emotion Research Group, Department of Psychology, University of Geneva, 2005
6. The role of emotion in a tractable architecture for situated cognizers [Текст] / Petta P. // In: Trappl R, Petta P, Payr S, editors. Emotions in humans and artifacts. Cambridge, MA: MIT Press; 2003
7. The Emotion Machine: Commonsense Thinking, Artifiial Intelligence, and the Future of the Human Mind [Текст]/ Minsky Marvin // Simon and Schuster, 2007
8. Temporal Sequence Learning, Prediction, and Control - A Review of different models and their relation to biological mechanisms [Текст] / Wörgötter F., Porr B. // Department of Psychology, University of Stirling, 2005
9. Affect and Proto-affect in effective functioning / Ortony A, Norman D, Revelle W. // In: Fellous J-M, Arbib MA, editors. Who need emotions? New York: Oxford University Press; 2005
10. The feeling of what happens: Body, emotion and the making of consciousness [Текст] / Damasio AR. // London: Vintage; 1999., LeDoux JE. The emotional brain. New York: Simon Schuster;  
    1996., Rolls E. Emotion explained. Oxford: Oxford University Press; 2005
11. Emotion and cognition: Insights from studies of the human amygdala [Текст] / Phelps E. // Annu Rev Psychol. 2006
12. On the nature and function of emotion: a component process approach / K. Scherer [Текст] // In: K.R. Scherer & P. Ekman Approaches to Emotion . Hillsdale, N.J.: Lawrence Erlbaum, 1984
13. Emotion and decision-making: affect-driven belief systems in anxiety and depression [Текст] / Martin P. Paulus1, Angela J. Yu // Trends in Cognitive Sciences September 2012, Vol. 16, No. 9
14. Prospect theory: an analysis of decision under risk [Текст]/ Kahneman, D., Tversky, A. // Econometrica 47, 263–291, 1979
15. A dual system model of preferences under risk [Текст] / Mukherjee, K. // Psychol. Rev. 117, 243–255; 2010
16. Music, pandas, and muggers: on the affective psychology of value [Текст] / Hsee, C.K. and Rottenstreich, Y. // J. Exp. Psychol. Gen. 133, 23–30; 2004
17. Preferences under risk: content-dependent behavior and psychological processing [Текст] / Kusev, P. and van Schaik, P. // Front. Psychol. 2, 269; 2011
18. Designing sociable robots [Текст] / Breazeal C. // Cambridge, MA: MIT Press; 2002
19. Neurochemical networks encoding emotion and motivation: An evolutionary perspective [Текст] / Kelley AE. // In: Fellous J-M, Arbib MA, editors. Who needs emotions? The brain meets the robot. New York: Oxford University Press; 2005
20. Neuromodulating Cognitive Architecture: Towards Biomimetic Emotional AI [Текст] / Max Talanov, Jordi Vallverdu, Salvatore Distefano, Manuel Mazzara, Radhakrishnan Delhibabu // Advanced Information Networking and Applications (AINA), 2015 IEEE 29th International Conference, ISSN : 1550-445X, Print ISBN: 978-1-4799-7904-2, 587 - 592
21. Психология: учеб. [Текст] / В. М. Аллахвердов, С. И. Богданова и др.; отв. ред. А. А. Крылов. // — М.: Проспект, 2005. С. 214—217
22. Artificial Cognitive Systems [Текст] / Vernon David // The MIT Press Cambridge, Massachusetts London, England: A Primer, 2014
23. Some philosophical problems from the standpoint of artificial intelligence at the Wayback Machine [Текст] / McCarthy, J., and Hayes, P. J. // (archived August 25, 2013). In Meltzer, B., and Michie, D., eds., Machine Intelligence, No. 4. Edinburgh: Edinburgh University Press, p. 463-502; 1969
24. Марвин Минский и эмоциональные машины [Электронный источник] / Таланов Максим // Режим доступа: https://postnauka.ru/faq/58727 – свободный
25. A new three-dimensional model for emotions and monoamine neurotransmitters [Текст] / H. Lövheim // Med Hypotheses. — 2012. — Vol. 8. — Pp. 341–348
26. Affect theory [Текст] / Tomkins S. // In: Ekman P, Friesen W, Ellsworth P, editors. Emotions in the human face. Cambridge: Cambridge University Press; 1982. p. 355–95
27. Emotion and Adaptation [Текст] / Smith, Craig A. & Lazarus, Richard S // In L.A. Pervin (Ed.) Handbook of Personality: Theory and Research. (pp. 609-637). New York: Guilford, 1990
28. Progress on a cognitive-motivational-relational theory of Emotion [Текст] / Lazarus, Richard S. // American Psychologist, 46(8), 819-834; 1991
29. Reinforcement learning in the brain [Текст] / Yael Niv // Psychology Department & Princeton Neuroscience Institute, Princeton University, 2009
30. Adaptive critic and the basal ganglia [Текст] / Barto, A. G. // In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), Models of information processing in the basal ganglia (p. 215-232). Cambridge: MIT Press., 1995
31. A neural substrate of prediction and reward [Текст] / Schultz, W., Dayan, P., & Montague, P. R // Science, 275, 1593-1599, 1997
32. Cellular models of reinforcement [Текст] / Wickens, J. R., & Kotter, R. // In J. C. Houk, J. L.Davis, & D. G. Beiser (Eds.), Models of information processing in the basal ganglia (p.187-214). MIT Press, 1995
33. Learning and sequential decision making [Текст] / Barto, A. G., Sutton, R. S., & Watkins, C. J. C. H. // In M. Gabriel & J. Moore (Eds.), Learning and computational neuroscience: Foundations of adaptive networks (p. 539-602). Cambridge, MA: MIT Press, 1990
34. Neuro-dynamic programming [Текст] / Bertsekas, D. P., & Tsitsiklis, J. N. // Athena Sc., 1996
35. Reinforcement Learning. An Introduction [Текст] / Sutton, R. S. and Barto, A. G. // Bradford Books, MIT Press, Cambridge, MA, 2002 edition, 1998
36. Dynamic Programming [Текст] / Bellman, R. E. // Princeton University Press, Princeton, NJ, 1957
37. Learning to predict by the methods of temporal differences [Текст] / Sutton, R. S. // Mach. Learn, 1998
38. Generalization in reinforcement learning: Successful examples using sparse coarse coding [Текст] / Sutton, R. S. // In Touretzky, D. S., Mozer, M. C., and Hasselmo, M. E., editors, Advances in Neural Information Processing Systems: Proceedings of the 1995 Conference, pages 1038–1044, Cambridge, MA, 1996
39. Problem solving with reinforcement learning. PhD thesis / Rummery, G. A. // Cambridge University, Cambridge, 1995
40. Learning from delayed rewards. PhD thesis [Текст] / Watkins, C. J. C. H. // University of Cambridge, Cambridge, England, 1989
41. Technical note: Q-Learning [Текст] / Watkins, C. J. C. H. and Dayan, P. // Mach. Learn., 8:279–292; 1992
42. Conditioned reflexes [Текст] / Pavlov, P. I. // Oxford University Press, London
43. Обучение с подкреплением (Reinforcement Learning) [Электронный источник] / Воронцов К.В. // Режим доступа: <http://www.machinelearning.ru/wiki/images/archive/3/35/20140621071329!Voron-ML-RL-slides.pdf> – свободный
44. A Markovian Decision Process [Текст] / R. Bellman // Journal of Mathematics and Mechanics 6, 1957
45. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement / Rescorla, R. A.,Wagner, A. R // In A. H. Black & W. F. Prokasy (Eds.), Classical conditioning II: Current research and theory (pp. 64–99). New York, NY: Appleton-Century-Crofts, 1972
46. NEST (NEural Simulation Tool) [Текст] / Gewaltig Marc-Oliver, Diesmann Markus // Scholarpedia. — 2007. — Vol. 2, no. 4. — P. 1430
47. Supercomputers Ready for Use as Discovery Machines for Neuroscience //  
    Frontiers in Neuroinformatics. — 2012. — november. — Vol. 6. — Pp. 1–12
48. Nest: an environment for neural systems simulations / Diesmann M Gewaltig M // 2001
49. Affective Computing / R.W. Picard // MIT Press, 1997.

# ПРИЛОЖЕНИЕ А

(справочное)

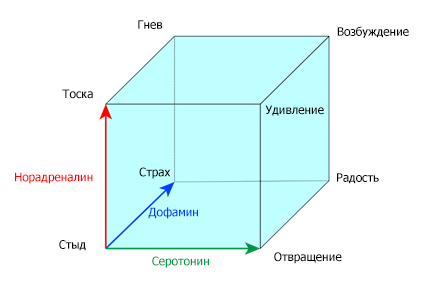
Куб эмоций Лёвхейма

Куб Лёвхейма – это теоретическая модель, описывающая связь между базовыми эмоциональными состояниями (страх, гнев, стыд, тоска, отвращение, удивление, радость, возбуждение) и влиянием трех нейромодуляторов (дофамин, серотонин, норадреналин).

Модель была предложена Хьюго Лёвхеймом в 2012 году. Куб эмоций помещен в ортогональную трехмерную систему координат, где лучами являются дофамин, серотонин и норадреналин. Вершины куба соответствуют низкому или высокому уровню того или иного нейромедиатора. Например, радости соответствуют высокое содержание дофамина и серотонина и низкое норадреналина, а страху – высокое содержание дофамина и низкое серотонина и норадреналина.

Куб Лёвхейма изображен на рисунке А.1.

Рисунок А1 – Куб Лёвхейма



# ПРИЛОЖЕНИЕ Б

(справочное)

Некоторые части естественного мозга

В этом приложении в алфавитном порядке описаны те части естественного мозга, которые непосредственно фигурируют в тексте.

|  |  |  |
| --- | --- | --- |
| Часть мозга | Английское название (сокращение) | Краткое описание |
| Бледный шар | Globus pallidus (GP) | Субкортикальная структура мозга, часть переднего отдела головного мозга, сохраняющая функциональные связи с гипотоламусом |
| Бледный шар (externa) | Globus pallidus pars externa (GPe) | Внешная часть бледного шара |
| Бледный шар (interna) | Globus pallidus pars interna (GPi) | Внутренняя часть бледного шара |
| Вентральная область покрышки | Ventral tegmental area (VTA) | Часть среднего мозга, начало дофаминовых путей. Играет роль в системах вознаграждения, а так же участвует в формировании зависимостей |
| Вентральное полосатое тело | Ventral striatum (VS) | Часть полосатого тела |
| Гипоталамус | Subtalamic nucleus (STN) | Отдел промежуточного мозга |
| Двигательная кора | Motor cortex (MC) | Часть коры головного мозга, участвующая в планировании, контроле и выполнении произвольных движений |
| Кора | Cortex (C) | Серое вещество, покрывающее полушария головного мозга |
| Палеостриатум | Ventral palidum (VP) | Серое вещество, расположеное в глубине каждого полушария головного мозга. Часть базальных ганглиев, включает в себя полосатое тело и чечевицообразное ядро |
| Полосатое тело | Striatum (S) | Часть базальных ганглиев, чередующиеся полосы белого и серого вещества. Включает в себя хвостатое ядро и чечевицообразное ядро |
| Префронтальная кора | Prefrontal cortex (PFC) | Отдел коры больших полушарий головного мозга, передняя часть лобных долей |
| Ретрорубальная область | Retrorubal area (RRA) | Часть ретикулярной формации среднего мозга |
| Черная субстанция (compacta) | Substantia nigra pars compacta (SNc) | Часть среднего мозга. Входит в состав экстрапирамидной системы. Принимает сигналы в цепи базальных ганглиев, поставляет дофамин полосатому телу |
| Черная субстанция (reticulata) | Substantia nigra pars reticulata (SNr) | Часть среднего мозга. Входит в состав экстрапирамидной системы. Передает сигналы от базальных ганглиев к различным частям мозга |