

**Федеральное государственное автономное образовательное учреждение высшего
профессионального образования**

КАЗАНСКИЙ (ПРИВОЛЖСКИЙ) ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ

**ВЫСШАЯ ШКОЛА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И
ИНФОРМАЦИОННЫХ СИСТЕМ**

Направление подготовки: 09.03.03 — Прикладная информатика

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

**ИНТЕГРАЦИЯ ТРЁХ СИСТЕМ НЕЙРОМЕДИАТОРОВ ДЛЯ
РЕАЛИЗАЦИИ БАЗОВЫХ ЭМОЦИЙ В КОГНИТИВНОЙ
АРХИТЕКТУРЕ NEUCOGAR**

Работа завершена:

«_____» _____ г.

Студент группы 11-304

_____ М. А. Загулова

Работа допущена к защите:

Научный руководитель

к. т. н., заместитель директора по науч-
ной деятельности

«_____» _____ г.

_____ М. О. Таланов

Директор Высшей школы ИТИС

«_____» _____ г.

_____ А. Ф. Хасьянов

Казань — 2017 г.

ОГЛАВЛЕНИЕ

| | |
|---|----|
| Введение | 3 |
| 1. Современное состояние исследований в области эмоциональных вычислений. | 6 |
| 2. Эмоциональные состояния на уровне нейробиологических процессов. | 9 |
| 3. Реализация норадреналиновой системы | 15 |
| 4. Интеграция норадреналиновой системы с серотониновой и дофаминовой системами. | 25 |
| 5. Симуляция базовых эмоций в трёхмерной модели | 33 |
| Заключение | 35 |

ВВЕДЕНИЕ

Идея построения машины, обладающей интеллектом, способной искать и принимать оригинальные решения, базируется на большем, чем математическая модель — на чувствах удовлетворения и горя, факторе внезапности, чтобы обучаться подобно ребёнку, писал А. Тьюринг. [small] [big] [intelligent_machinery]] Связь мышления с эмоциями подтверждает и объясняет М. Мински в труде об искусственном интеллекте и эмоциональном интеллекте. [minsky2007]

Эмоция, по соглашению психологов и нейробиологов, это адаптивное согласованное изменение в нескольких физиологических системах (соматических и нервных) в ответ на внешний или внутренний стимул.[plutchik2001] Большую часть прошлого века разум и эмоции было принято противопоставлять. На настоящее же время нейробиологами многократно доказано, что машинное понимание, рациональное, логическое мышление и эмоции тесно переплетены и неразделимы.[Brosch; ziemke-role; damasio2004; damasio2008]

Эволюционно эмоции обеспечивают выживание, предупреждая организм об угрозах, регулируют поведение, направляя организм на удовлетворение актуальных потребностей.[whoneedsemotions] Эмоции контролируют внутреннее психологическое состояние, почти всегда управляют принятием решений.[roleofemotion] Эмоции участвуют в процессах обучения, памяти, управляют вниманием и восприятием, механизмы сознания и эмоций связаны на всех уровнях.[Phelps2006] Социальные взаимодействия реальных людей с субъектом без эмоций затруднены и непродуктивны, поскольку эмоции регулируют адекватность в общении, вычисляют наиболее подходящий к ситуации и настроению собеседника ответ. Эмпатия, жалость и любовь это сложные социальные эмоции, без наличия которых взаимодействие с людьми просто потенциально опасно. Эмоции нужны для гибкости, ускоренной адаптации и развития. Исходя из этих фактов, во-первых, понимание человеческих эмоций это необходимый шаг для понимания человеческого сознания. Во-вторых, искусственный интеллект и когнитивные технологии неосуществимы без реализации эмоций машины, вычислительных эмоций.[affectivecomputing]

Исследования по воспроизведению эмоций в вычислительных машинах называются Affective Computing. [affectivecomputingchallenges] При наличии различных подходов к этой проблеме, на данный момент не существует таких, которые бы опустились по уровню абстракции до уровня нейробиологических процессов в головном мозге — при этом сама идея логична, ведь именно такой уровень мыслительных процессов теоретически сопоставим с вычислительными процессами. Моделированием на этом уровне занимаются в лаборатории машинного понимания Высшей школы ИТИС, где разрабатывается когнитивная архитектура NeuCogAr (Neural Cognitive Architecture). Данная работа — завершение двухлетнего проекта по реализации трёх систем нейромедиаторов для воспроизведения 8 базовых эмоций.

Объектом исследования в дипломной работе являются вычислительные эмоции в биологически реалистичной когнитивной архитектуре. Предмет исследования — принципы влияния нейромедиаторов норадреналин, серотонин, дофамин, ацетилхолин, ГАМК и глутамат на центральную нервную систему и друг на друга; перенос этих механизмов на область вычислительных технологий.

Цель дипломной работы — интегрировать три системы нейромодуляторов в когнитивной архитектуре NeuCogAr для реализации 8 эмоциональных состояний на биологически реалистичном количестве нейронов. Для достижения поставленной цели необходимо решить следующие задачи:

- Реализовать систему распространения нейромодулятора норадреналин на NEST Simulator (инструменте моделирования импульсных нейронных сетей);
- Совместить программные реализации систем норадреналина, серотонина и дофамина;
- Исследовать воздействие трёх моноаминовых нейромодуляторов друг на друга в мозге млекопитающего;
- Воспроизвести это воздействие на вычислительной нейронной сети в 500 000 нейронов (биологически реалистичное количество для мыши);

- Запустить полученную систему в разных конфигурациях, согласно модели Хьюго Левхейма, фиксируя 8 различных эмоций машины.

Основным инструментом в дипломной работе является симуляционное ядро NEST Initiative, находящееся под контролем пакета языка Python. Его основным пользовательским интерфейсом является PyNEST, который удобен в использовании и хорошо взаимодействует с библиотеками языка Python, например, matplotlib, numpy, pylab.

ГЛАВА 1. СОВРЕМЕННОЕ СОСТОЯНИЕ ИССЛЕДОВАНИЙ В ОБЛАСТИ ЭМОЦИОНАЛЬНЫХ ВЫЧИСЛЕНИЙ.

Современные социальные роботы могут имитировать эмоции в соответствии с выражением лица собеседника: андроид София из Hanson Robotics показывает 60 эмоций; андроид ERICA из Hiroshi Ishiguro Laboratories имитирует выражение лица собеседника, если в её сценарии нет ответной эмоции на его реплику [erica]; социальный робот Pepper распознаёт эмоцию собеседника по выражению лица, тону голоса и ключевым словам, реагирует согласно заранее заданным сценариям. [meetpepper] Все они известны именно как «эмоциональные», обладающие «эмоциональным движком», создают иллюзию наличия эмпатии, но не являются когнитивными технологиями и не испытывают эмоций.

Иной подход к эмоциям показан С. Бризейл: эмоции как система мотивации, эмоции как реакция на то, достигнуты ли цели и стремления организма. [Breazeal] Её робот KISMET обладает стремлениями общаться с людьми, наблюдать за яркими игрушками и периодически отдыхать — как человеческий младенец. Если жизненные цели робота не выполняются, то в нём нарастают негативные эмоции (грусть, раздражение) как средство мотивации к их выполнению. По достижению цели робот испытывает положительные эмоции, которые служат сигналом успокоиться и отдохнуть. Если же стимуляция робота перевыполнена, к нему подошли слишком близко, пугают его — эмоции робота вновь станут негативными (страх), мотивирующими избежать нежелательных стимулов, например, отвернуться, закрыть глаза.

Модель эмоций, мотивирующих робота KISMET, опирается на «колесо эмоций», модель, созданную Р. Плутчиком, группирующую в пары 8 базовых эмоций: грусть — радость, страх — ярость, отвращение — приятие, удивление — разочарование. Процесс влияния эмоций на организм Плутчик описывает как последовательность событий: стимулирующее событие — ожидание реакции — ощущение эмоции — физиологическое возбуждение — импульс к действию — совершение действия — влияние действия. [plutchik2001] С учётом этой модели в 2012 году Э. Камбрия и др. разработали биологически-

инспирированную и опирающуюся на психологию модель представления любой эмоции в пространстве 4 измерений: внимание, чувствительность, удовольствие, уверенность.[**hourglass**]

На междисциплинарном симпозиуме AAAI по эмоциональным архитектурам были представлены взгляды на моделирование эмоций с различных сторон. Физиологическая модель с максимально упрощёнными частями организма: дыхательной системой, сердцем и пр.[**aaai1**] Многоуровневая модель, учитывающая логические рассуждения, убеждения и систему Брамса о человеческом поведении.[**aaai2**] Простые эмоции для выстраивания победной тактики роботической рукой в игре «камень — ножницы — бумага», учитывающие удовольствие от победы и злость от поражения.[**aaai3**] Моделирование эмоций игрока обучающей игры с помощью динамических байесовских сетей.[**aaai4**] Эмоциональная архитектура для робота Казимира на основе BDI (архитектуры убеждений, желаний и намерений).[**aaai5**] Эмоциональная архитектура на основе парадигмы принятия решений по нескольким критериям с использованием нечёткой логики.[**aaai6**]

Когнитивист Т. Цимке не отделяет эмоции от всех процессов познания в целом, его эмоционально-когнитивные архитектуры всегда базируются на нейрофизиологии, моделировании участков мозга в соответствии с реальными биологическими функциями. Например, робот-крыса с чувствительными усиками для проекта ICEA имеет мозжечок, гиппокамп, кору мозга, базальные ядра, и каждый этот участок имеет свою функцию, связанную с памятью либо мотивацией. В этих архитектурах эмоциональные состояния несут конкретные функции для выживания, поддержания активности, памяти, предсказаний и планирования.[**ziemkeflexibility**]

Отдельно идёт развитие автономных военных роботов, которых исследователи учат принимать решения, связанные с моралью и этикой, в вопросах доверия, жертвенности и вины.[**moralrobots**] Речь также идёт об эмоциях, классифицированных Дж. Хальдтом как «моральных»: уважение, благодарность, сострадание, презрение, стыд, вина и ярость.[**moralemo**] Они реализованы не биологически реалистично, а подсчётом сложных алгоритмов, принимающих

решение поступенчато, не исключая и законов Азимова.[**militaryrobotics**]

ГЛАВА 2. ЭМОЦИОНАЛЬНЫЕ СОСТОЯНИЯ НА УРОВНЕ НЕЙРОБИОЛОГИЧЕСКИХ ПРОЦЕССОВ.

Исследование эмоциональных вычислений как поиск моделей, объясняющих эмоции, начиналось в областях психологии, нейропсихологии и философии. Эволюционная теория происхождения эмоций Ч. Дарвина объясняла эмоции как приспособительные механизмы, жизненно важные для адаптации организма к условиям жизни — неважно, человека или животного. Д. Пейпекс выдвигал гипотезу о единой системе, объединяющей структуры мозга, участвующие в формировании эмоций, — позже она получила название лимбической системы. Исследования Ф. Барда, подкрепленные данными физиологии, свидетельствовали, что физиологические проявления и субъективные переживания эмоциональных процессов происходят одновременно, благодаря расщеплению возбуждающего нервного импульса в таламусе. Двухфакторная теория эмоций С. Шехтера разделяла эмоцию на эти же два компонента: физиологическое возбуждение и его когнитивную интерпретацию.

Десятилетиями нейробиологи публиковали результаты экспериментов о том, нейроны какого типа и в какой части мозга проявляют активность, в зависимости от состояния организма, в том числе и эмоционального состояния. Эти многочисленные данные собраны и обобщены доктором психиатрии Хьюго Левхеймом, который предложил на их основе новую трёхмерную модель эмоций. [lovheim2012] Базовые эмоции (аффекты), объясняемые в его работе нейробиологически, берутся из психологической модели Сильвана Томкинса. [tomkins1963; tomkins1991] Теория аффектов Томкинса разделяет 9 «первичных» эмоций (наблюдаемых с рождения) на категории и соединяет каждую с типичным физиологическим проявлением, позой, выражением лица. Например, злость — покрасневшее лицо, сжатые челюсти, сдвинутые брови. Из этих 9 эмоций Левхейм не стал вводить в модель эмоцию реакции на неприятный запах, остальные 8 включены и присутствуют на рисунке 1.

Тремя измерениями модели Левхейма являются уровни концентрации моноаминовых нейромедиаторов серотонин, дофамин и норадреналин — минимально необходимое количество для того, чтобы специально упрощённая мо-

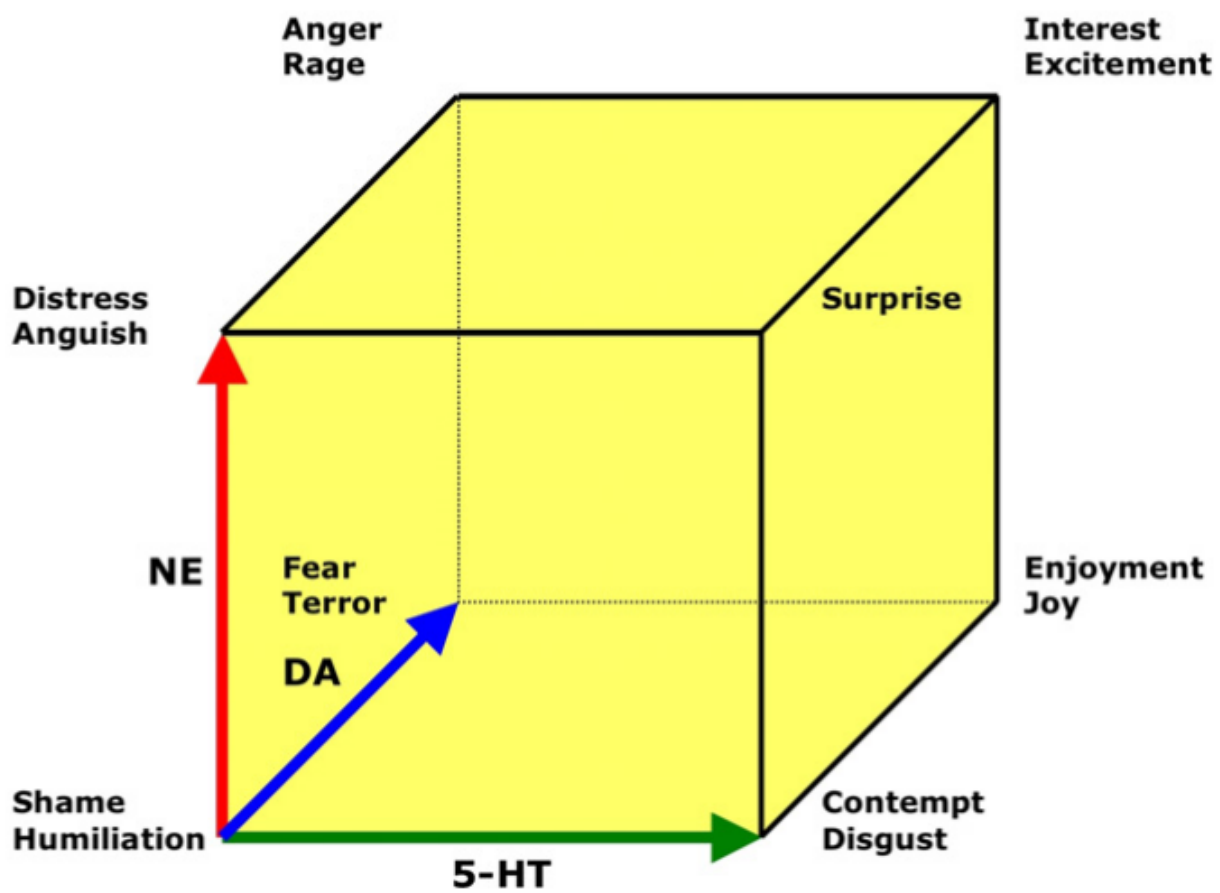


Рисунок 1. Модель Левхейма, задающая трёхмерное пространство эмоций через нейромедиаторы серотонин, дофамин и норадреналин.

дель оставалась реалистичной. Нейромедиаторы это биологически активные химические вещества, контролирующие процессы, происходящие между нейронами. Системы названных моноаминовых нейромедиаторов эволюционно сохраняются в различных организмах, они контролируют поведение как людей, так и крыс, лобстеров, пустынной саранчи и др. Эти системы обеспечивают гибкость поведения и способность адаптироваться к постоянно изменяющейся окружающей среде, сами по себе также являются динамичными. [Benarroch1711200]

Нейронная активность может модифицироваться и перенастраиваться, когда концентрация нейромедиатора повышается из-за активности конкретных зон мозга (например, голубоватого пятна, в случае норадреналина). Нейромедиатор разносится по другим областям мозга, где влияет на многочисленные рецепторы подходящего типа. Это может усиливать или ослаблять связь между нейронами, если связь также была активна совсем недавно (эффект собаки Павлова), менять свойства нейронов. [parsingreward] У таких нейромодулирующих систем существуют анатомические привилегии широкого доступа к многим областям нервной системы, и они идеально расположены для регулирования всех процессов обработки и хранения информации в мозге. Нейромодулирующие системы принимают активное участие в работе двигательных, чувственных и когнитивных функций. [zarrindast] Их роль в работе мозга настолько велика, что, несмотря на внушительное количество уже имеющихся экспериментальных данных, вопросы взаимодействия нейромедиаторов друг с другом, точного механизма их функционирования и прочие остаются открытыми, требующими длительного дальнейшего изучения.

На рисунке 1 зелёная ось «5-НТ» обозначает уровень концентрации нейромедиатора серотонин. В человеческом организме серотонин отвечает за регуляцию поведения, аппетит, обучение, общение — за всё, что делают в спокойном состоянии. Низкий уровень серотонина влечёт сонливость, высокий уровень серотонина повлечёт за собой половое влечение, быстрое переваривание пищи, чувство эйфории и ощущение непобедимости. Серотонин работает над уверенностью в себе, особенно после верно принятых решений, и над контро-

лем агрессии. В точке, где концентрация серотонина максимальна, а остальных минимальна, расположена эмоция сытости/отвращения/презрения.

Красная ось «NE» - уровень концентрации нейромедиатора норадреналин, или норэпинефрин. Уровень его концентрации повышается в ответ на стрессовые ситуации, раздражители и на всё новое, либо неожиданное. Высокий уровень деятельности норадреналина вызывает повышение бдительности и скорости реакции, концентрацию внимания, улучшение обработки сенсорных входов, повышение уровня возбуждения. В точке, где концентрация норадреналина максимальна, а остальных минимальна, расположена эмоция страдания/отчаяния. Мозг испытывает ощущение бедствия, когда произошло что-то неоправимое или болезненное, и максимально сконцентрирован на этом событии.

Синяя ось «DA» - уровень концентрации нейромедиатора дофамин. Он отвечает за наслаждение и удовольствие от выполненного дела, вкусной еды, приятных ощущений, даёт ощущение награды, таким образом закрепляя важные для выживания и благосостояния действия. Дофамин используется мозгом для оценки действий и мотивации, а также для активации моторных функций. В точке, где его концентрация максимальна, а остальных минимальна, находится эмоция «страх».

В точке, где концентрация каждого нейромедиатора минимальна, располагается эмоция «стыд», когда нет ни сфокусированности внимания, ни чувства удовлетворения, ни уверенности в себе. Повышенные уровни дофамина и норадреналина влекут «ярость», серотонина и норадреналина - «удивление», дофамина и серотонина — спокойное «удовольствие». Высокий уровень концентрации всех трёх нейромедиаторов в модели Левхейма связан с эмоцией активной «заинтересованности», «увлеченности».

Отображение клеточных реакций на вычислительный процесс должно быть осуществлено так, чтобы при удалении всей биохимии, отсутствии настоящих моноаминов и белков удалось сохранить роли нейромедиаторов в системе. Были предложены следующие аналогии [talanov2014], зафиксированные на рисунке 2:

- Серотонин регулирует вычислительную мощность и выделяет постоянную память для хранения данных. Поскольку серотонин участвует в обучении и накоплении опыта, то результаты принятых решений, во-первых, нужно запомнить в качестве удовлетворительного или стыдного опыта (выделить для них память). Во-вторых, в зависимости от результата, за решением следует прилив сил и уверенности (мощности) или их упадок.
- Норадреналин отвечает за перераспределение вычислительных процессов и выделения оперативной памяти. Это соответствует переключению внимания в организме живого существа: мыслительные процессы тоже перераспределяются, чтобы полностью сконцентрироваться на том, что привлекло внимание.
- Дофамин, во-первых, выделяет постоянную память для хранения данных, как и серотонин — поскольку тоже участвует в обучении. Во-вторых, поскольку дофамин сильно влияет на мотивацию, ??????????????????????????????

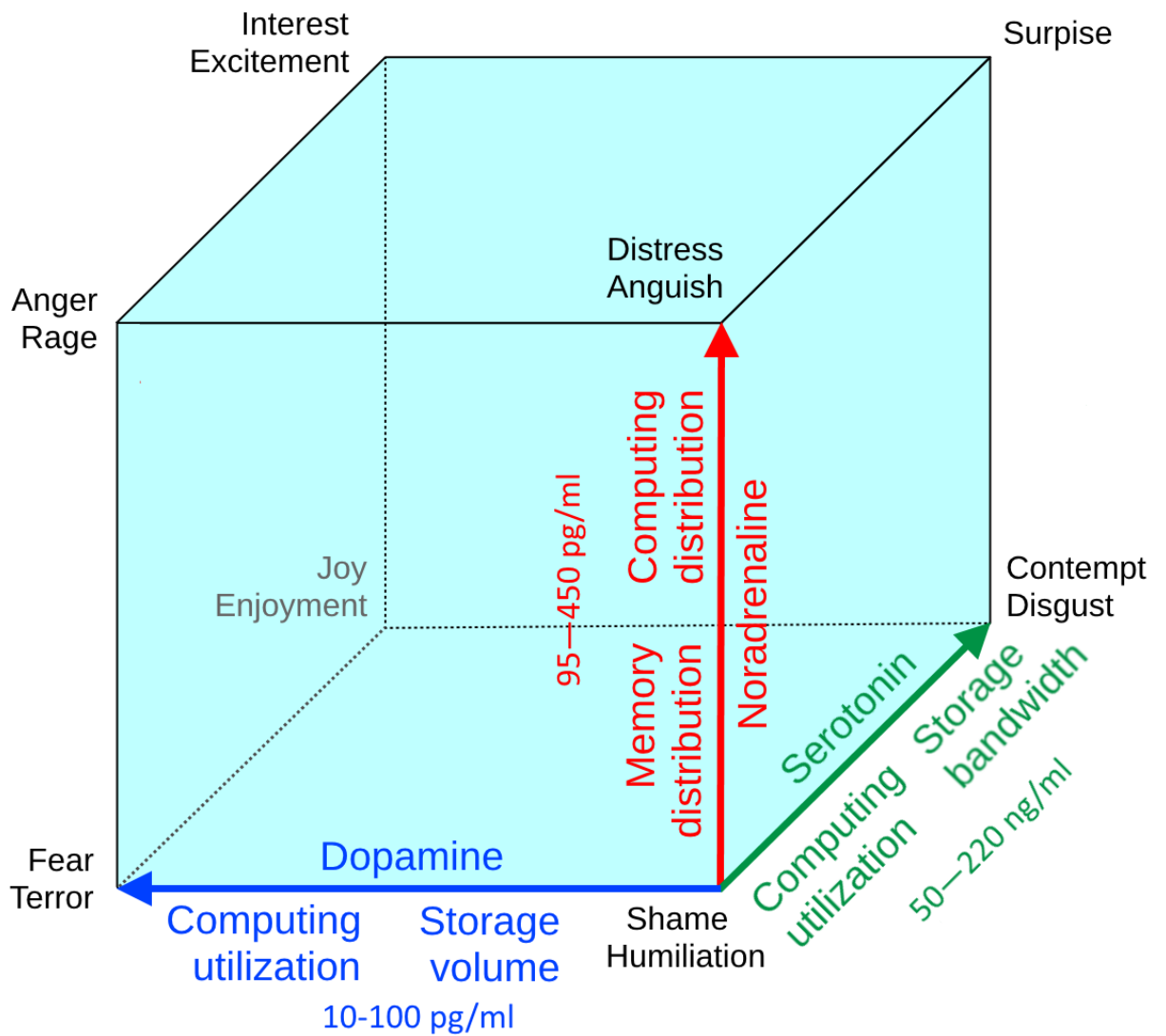


Рисунок 2. Дополненная модель Левхейма. Проведение аналогии роли нейромедиаторов с вычислительными процессами.

ГЛАВА 3. РЕАЛИЗАЦИЯ НОРАДРЕНАЛИНОВОЙ СИСТЕМЫ

Норадреналиновая система являлась последним недостающим элементом когнитивной архитектуры NeuCogAr, при уже реализованных дофаминовой и серотониновой системах. В этой главе описываются базовые принципы динамики концентрации норадреналина, их программная реализация и тестирование.

Нейромедиатор норадреналин (он же норэпинефрин) относится к биогенным аминам, группе катехоламинов, он активно используется симпатической периферийной нервной системой (мобилизует организм к действию) и центральной нервной системой (мобилизует мозг для действий).[**noradrenalin3**] Норадренергические нейроны в мозге немногочисленны: около 4000 в основном месте залегания — голубоватом пятне (locus coeruleus), около 3000 в ядрах солитарного тракта (nucleus tractus solitarii).[**masuko; rat_data5**] Но эти нейроны посылают проекции во многие области мозга и оказывают на них значительное влияние. Активность норадренергических нейронов связана с разнообразными реакциями: на стресс, на неожиданность, на такие раздражители как боль, жар или холод, затруднение дыхания. Во время сна активность в очаге голубоватого пятна падает, во время бодрствования работает на уровне минимально необходимом, резко повышается при появлении стимула, привлекающего внимание.[**Berridge2003**] Высокий уровень деятельности норадреналина вызывает повышение бдительности и скорости реакции, фокусировку внимания, улучшение обработки сенсорных входов, повышение уровня возбуждения.

Контроль над такими когнитивными процессами как внимание, возбуждение и бдительность необходим для реализации эмоциональных состояний злости, удивления, страдания и заинтересованности. Поэтому, по аналогии с существующими моделями дофамина и серотонина, принципы работы норадреналина были реализованы в виде нового модуля, добавленного в NEST Simulator.

Модель механизма изменения концентрации норадреналина основана на модели, предложенной Е. Ижикевичем для описания динамики концентрации дофамина после серии событий (поощрений).[**izhikevich**] В работе Е. Ижике-

вича принципы всплеска дофамина в ответ на поощрение иллюстрируются собаками Павлова, вырабатывавшими рефлекс в связи с тем, какие действия сопровождалась наградой. Исследуя механизм вырабатываемого рефлекса, неясно, как мозг «запоминает» цепочки нейронов, активность которых вознаграждалась - ведь между событием и наградой проходило несколько секунд (слишком долго для стандартного алгоритма спайковой пластичности).

Е. Ижикевич вводит понятие «сигнализирующей отметки» (eligibility trace) — своеобразных следов, оставляемых в синапсе после того, как по нему проходит спайк. В реальном биологическом мозге это какие-либо кислоты, активизируемые ненадолго, но дольше, чем на пару секунд. Затем, как проиллюстрировано на 3, награда (поощрение) вызывает резкое повышение уровня дофамина, и тогда, благодаря совпадению всплеска дофамина с «сигнализирующей отметкой» в синапсе, синаптическая связь усиливается (а в модели увеличивается её «вес»).

Сохранено предположение, что в синапсах, соединяющих два возбуждающих нейрона, действует спайковая пластичность, зависящая от времени. Модель Е. Ижикевича для дофамина была представлена в виде набора определяющих уравнений. Опираясь на них, для сети нейронов, где 80% возбуждающие и 20% ингибирующие, для норадреналина составлены следующие уравнения:

1) Динамика каждого нейрона такова, что потенциал мембраны v каждого нейрона в каждый момент времени зависит от абстрактной переменной восстановления мембраны u (новый потенциал в этой точке \dot{v}) [**tactile**]:

$$\dot{v} = k(v - v_{rest})(v - v_{thresh}) - u + I \quad (1)$$

$$\dot{u} = a * b * (v - v_{rest}) - u \quad (2)$$

$$if(v \geq 30[mV]) : \{v = -65[mV], u = u + 2[mV]\} \quad (3)$$

В нашей модели максимальный потенциал мембраны v_{thresh} и потенциал покоя v_{rest} являются константами, синаптический ток I (внутри нейрона) име-

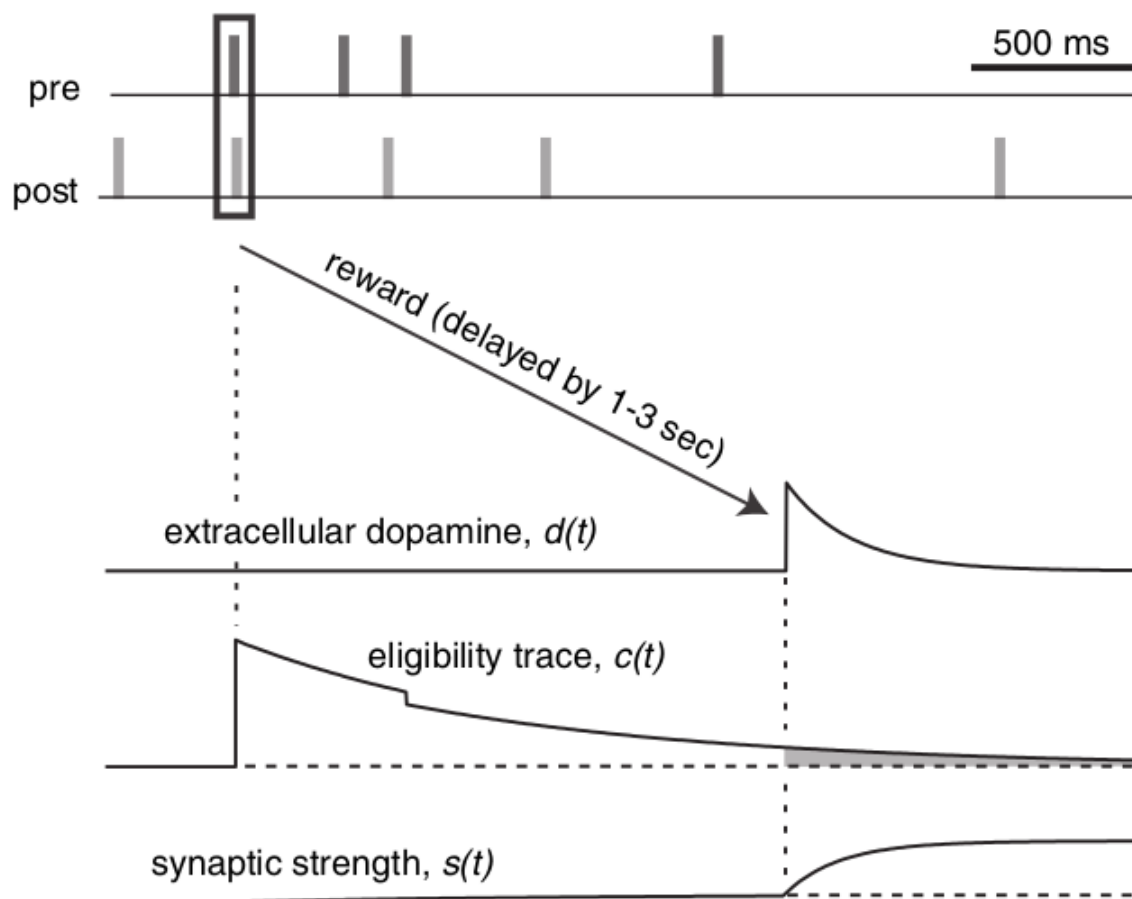


Рисунок 3. Иллюстрация процесса усиления синаптической связи из работы [izhikevich]. Происходит благодаря повышению уровня дофамина и имеющейся у синапса «сигнальной отметки» (eligibility trace), которая возникла секундами раньше.

ет экспоненциальную форму. Спайк происходит, когда потенциал мембраны превышает -50 mV , и после спайка потенциал мембраны восстанавливается: v снижается до -65 mV , u вырастает до 2 mV . Значение переменной a установлено 0.02 , переменная b равна 0.2 , коэффициент k — 1 .

Синаптический вес w (сила синаптической связи) изменяется в модели не напрямую [stdp], а модулируется через временную «сигнальную отметку» c , что показано в (6). Изменение «сигнальной отметки» c в течение времени происходит по следующему закону (новое значение в конкретной точке — \dot{c}):

$$\dot{c} = -\frac{c}{\tau_c} + A^+ e^{\frac{(t_{pre}-t_{post})}{\tau^+}} \delta(t - t_{post}) - A^- e^{\frac{(t_{pre}-t_{post})}{\tau^-}} \delta(t - t_{pre}) \quad (4)$$

где t_{pre} и t_{post} это время пресинаптического и постсинаптического спайков, A^+ и A^- это амплитуды изменения синаптического веса, τ_+ и τ_- — константы затухания, $\delta(t)$ это дельта-функция Дирака, пошагово увеличивающая переменную c . Значение сигнальной отметки уменьшается с постоянной скоростью τ_c .

Концентрация норадреналина также влияет на модулирование синаптического веса [nora1][nora2], что показано в (6). Концентрация норадреналина n уменьшается экспоненциально с течением времени (со скоростью τ_n), а увеличивается в зависимости от неожиданных событий:

$$\dot{n} = -\frac{n}{\tau_n} + p_{nov}n(\delta(t - t_n)p_{rew} + \delta(t - t_n)p_{pun}) \quad (5)$$

где p_{punish} это стрессовое событие (наказание), p_{rew} это событие награды, p_{nov} это вероятность того, что событие является новым и неожиданным. Концентрация норадреналина не может опуститься ниже нуля, и повышается из-за неожиданных событий как позитивной, так и негативной окраски.

Синаптический вес w (сила синаптической связи) изменяется в модели пропорционально относительной концентрации норадреналина n по сравнению с минимальной концентрацией (b_n), умноженной на значение сигнальной отметки c :

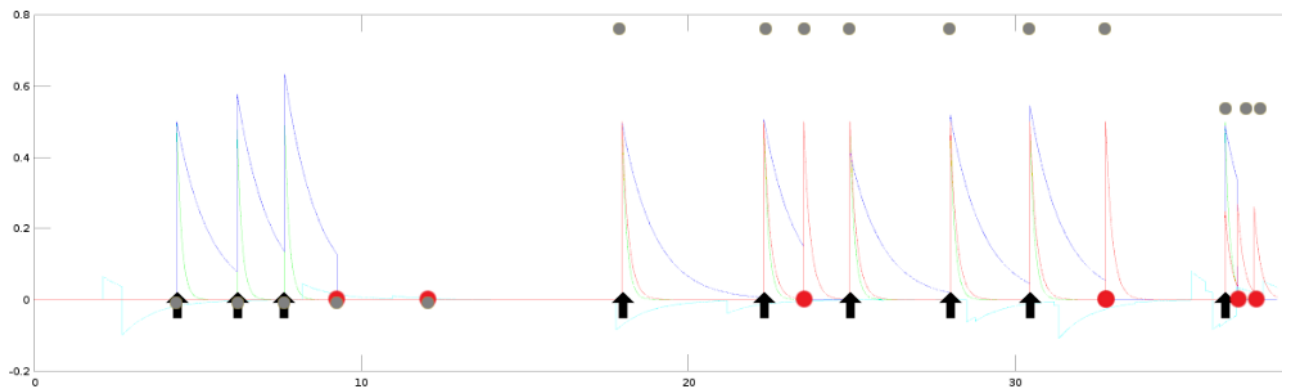


Рисунок 4. Изменение концентрации норадреналина (красный), дофамина (зелёный) и серотонина (синий) в зависимости от примененных к ним стимулов: поощрение (чёрная стрелка), стресс (красный кружок), новизна (серый кружок).

$$\dot{w} = c(n - b_n) \quad (6)$$

Модель была протестирована на языке MATLAB, вместе с уже реализованными моделями серотонина и дофамина, в следующих условиях:

- Задана сеть из 1000 нейронов с синаптической пластичностью, зависящей от времени;
- 100 синапсов на нейрон;
- Задержка проводимости 1 мс;
- Максимальный вес синапса 5;
- Потенциал покоя мембраны -65 mV;

На ?? изображён результат работы нейронной сети из 1000 нейронов в течение 500 миллисекунд, на которую подавались случайным образом события поощрения и стресса. Каждое из событий имело уровень неожиданности нулевой (полностью предсказуемое событие, первая треть временной шкалы), средний (последняя треть временной шкалы) или высокий (резкое, внезапное событие). При абсолютной предсказуемости событий концентрация норадреналина не поднимается совсем, тогда как уровни серотонина и дофамина одинаково поднимаются от предсказуемой награды, от предсказуемого наказания

(стресса) уровень серотонина обрушивается вниз, а дофамин на наказание не реагирует. При событии резком, сколько-нибудь внезапном, будь то поощрение или наказание, концентрация норадреналина повышается; реакция серотонина и дофамина на эти события остаётся той же, что и когда они были «ожидаемы», «предсказаны» системой.

Полученные реакции соответствуют экспериментальным результатам исследований работ разных лет. [viljoen; delaneytransmission; macaca] Поскольку результат, показанный моделью норадреналина в MATLAB, был убедителен, то модель была перенесена на язык c++ и добавлена в инструмент симулирования нейронных спайковых сетей NEST-2.12.0 в качестве нового модуля. Работа модуля была протестирована на простой модели из трёх нейронов: пресинаптический нейрон, постсинаптический нейрон и норадренергический нейрон. Модель симулировала процесс нейромодуляции, усиливающий связь между обычными нейронами в случае, если между ними произошёл спайк при высокой концентрации нейромедиатора норадреналин. Для этого к каждому из трёх нейронов были подключены генераторы спайков в 250 Гц и детекторы спайков. Модель была запущена дважды: с активацией норадреналина и без. Симуляции длились по 3000 миллисекунд.

На рисунке 5 показана спайковая активность каждого из трёх нейронов в эксперименте с активированным норадреналином. На рисунке 6 — изменение концентрации норадреналина со временем. Сильное увеличение концентрации наступает после относительно «длительного» отсутствия спайков, поскольку после «тишины» в 400 миллисекунд новый спайк имеет уже ненулевой фактор внезапности, считается неожиданным. В случае же спайков идущих подряд концентрация падает. То, как сочетание спайков и изменений концентрации повлияло на силу связи (вес синапса) между пресинаптическим и постсинаптическим нейронами, показывает рисунок 7. Вес синапса растёт рывками, в соответствии с работой генераторов.

Второй, контрольный эксперимент без активации норадреналина показал, что без присутствия нейромедиатора сила связи между нейронами единична и не меняется на протяжении всей симуляции (рисунок 8).

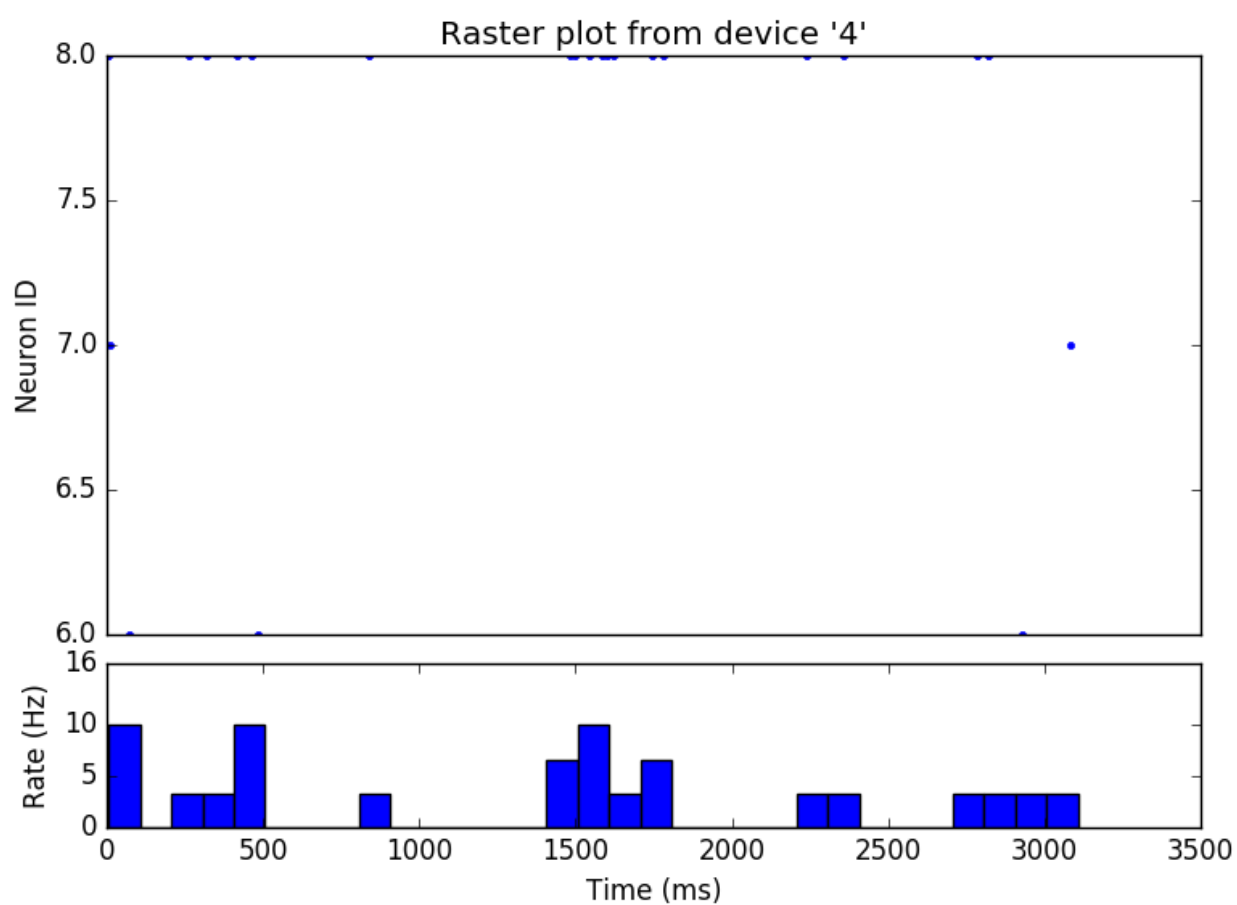


Рисунок 5. Спайковая активность на нейроне 6 (пресинаптический), нейроне 7 (постсинаптический), нейроне 8 (норадренергический).

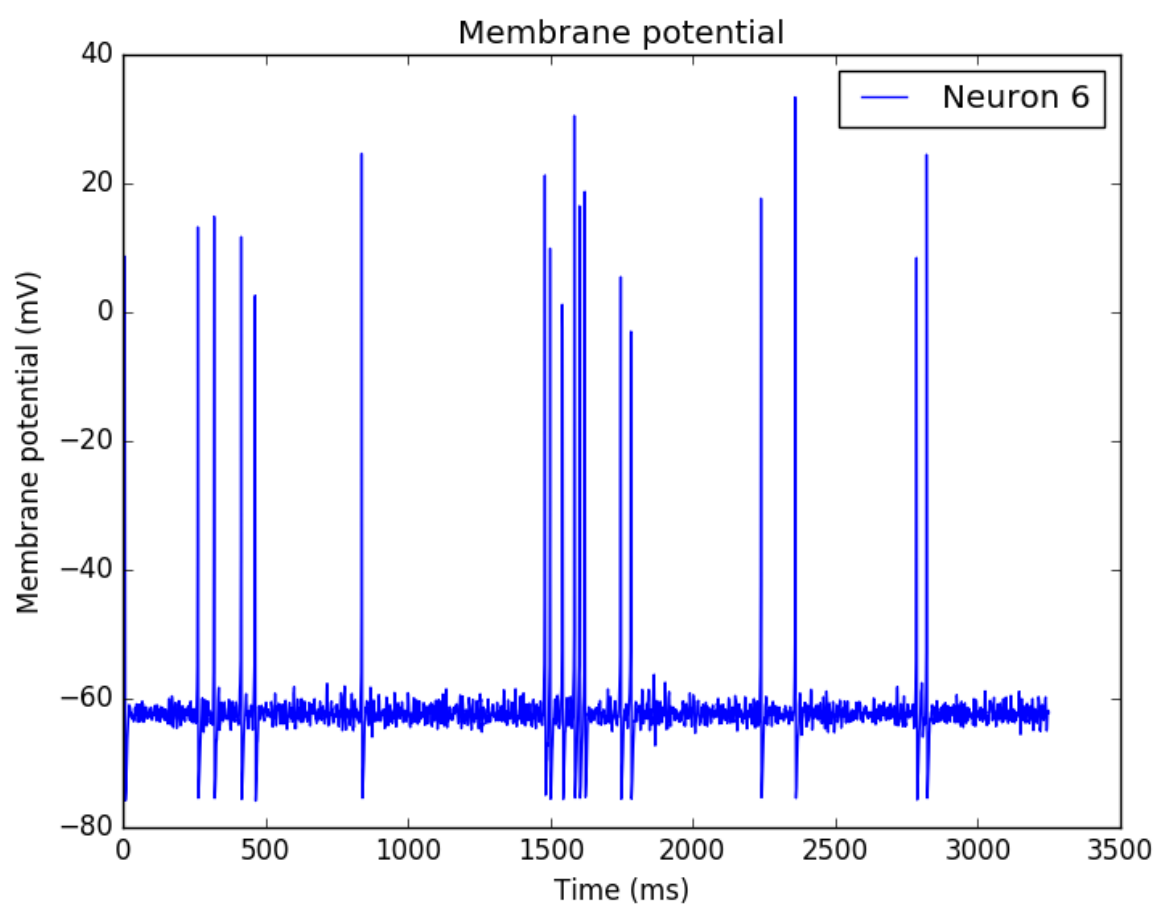


Рисунок 6. Динамика концентрации норадреналина.

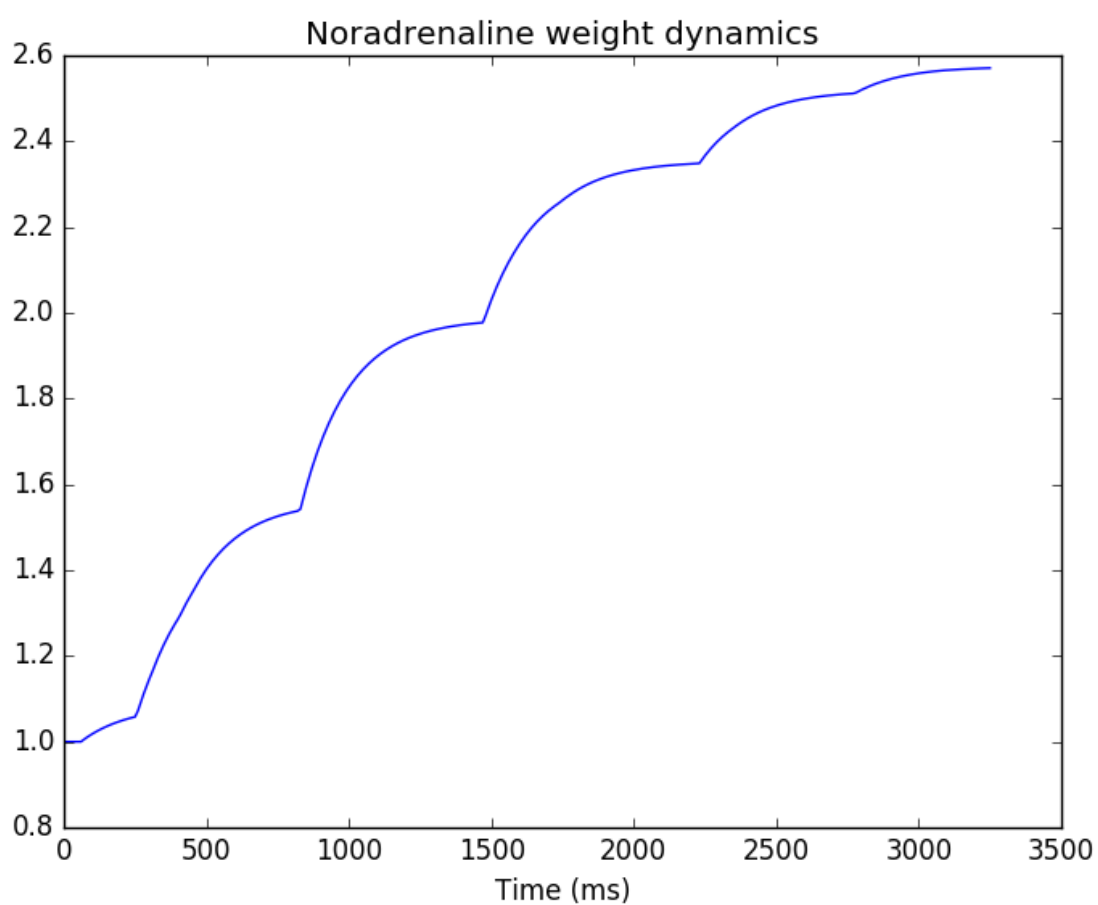


Рисунок 7. Изменение силы связи между пресинаптическим и постсинаптическим нейронами в присутствии норадреналина.

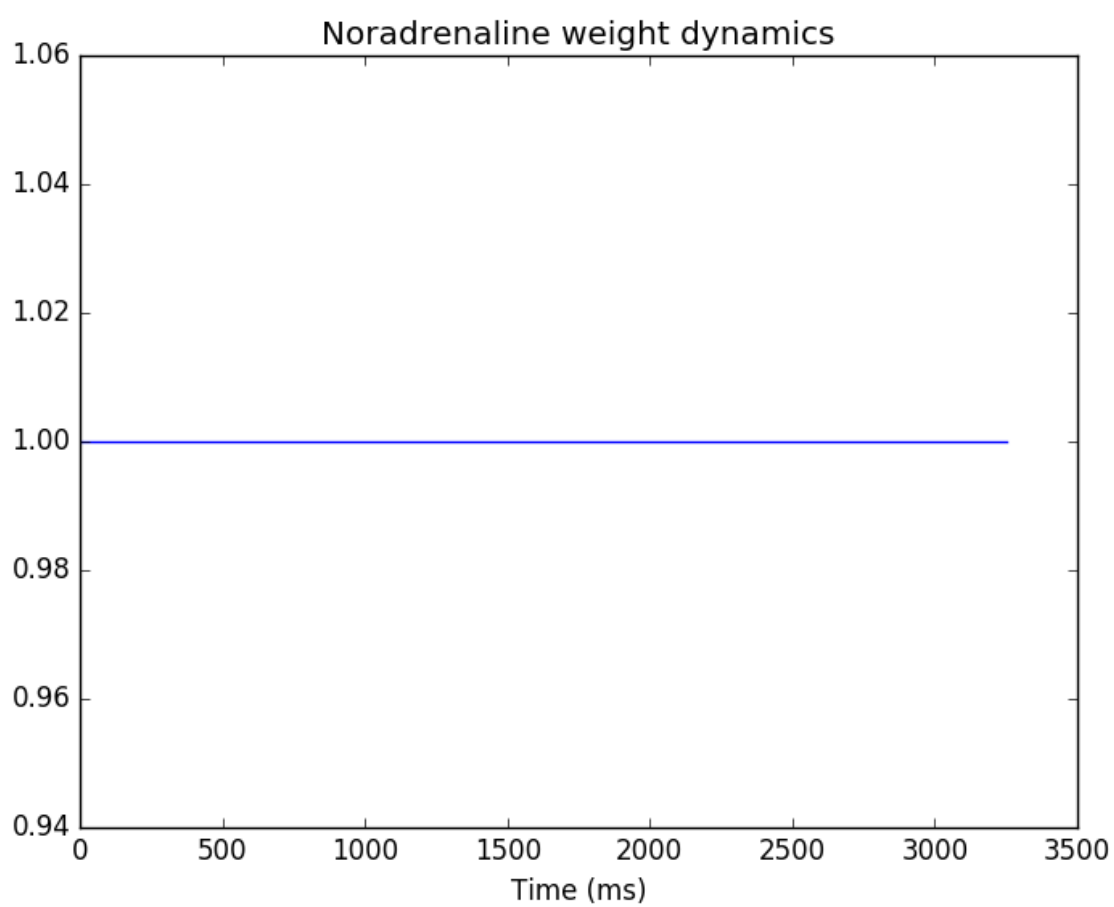


Рисунок 8. Изменение силы связи между пресинаптическим и постсинаптическим нейронами в отсутствие нейромедиаторов.

ГЛАВА 4. ИНТЕГРАЦИЯ НОРАДРЕНАЛИНОВОЙ СИСТЕМЫ С СЕРОТОНИНОВОЙ И ДОФАМИНОВОЙ СИСТЕМАМИ.

Схема распространения норадреналина по мозгу была схематично изображена выпускником ИТИС 2016 года Д. Седленко. Эта схема представлена на рис. 9:

Схема была реализована на языке python для инструмента моделирования нейронных спайковых сетей NEST-2.12.0. Каждой области мозга, указанной на схеме, было выделено биологически реалистичное для мозга крысы количество нейронов (данные о количестве нейронов собраны выпускницей ИТИС 2016 года Ю. Сафандеевой). Области мозга, в свою очередь, разбивались на подобласти, взаимодействующие с разными нейромедиаторами и друг с другом: например, ядрах шва есть два вида рецепторов норадреналина (на них выделено по 2900 нейронов) и серотониновые нейроны. Для активации источников норадреналина генераторы спайков должны быть подключены к зоне ядра одиночного пути, латерально-спинному ядру покрышки, вентральной области покрышки, перифинальной коре и паравентиккулярному ядру. Существуют аналогичные программные реализации систем дофамина (схема на рис. 10) и серотонина (схема в общем доступе по ссылке <https://github.com/research-team/NEUCOGAR/NES>). Различные области мозга, затрагиваемые нейромедиаторами на трёх схемах, были собраны в один словарь на python, количество нейронов на каждой – сохранено. Нейромодулирующие связи были последовательно подключены, и без добавления новых связей три схемы оказались соединены через вентральную область покрышки, префронтальную и моторную кору, ядра шва, миндалевидное и полосатое тело. Норадреналин, по схеме, должен возбуждать таламо-кортикальный цикл, в котором работает дофамин. Серотонин должен ингибировать центры дофамина, но возбуждать префронтальную кору, миндалевидное тело и гиппокамп. Норадреналин вызывает активность в ядрах шва, которые содержат центры серотонина, но последующее преодоление норадреналином ингибирующего влияния серотонина спорно. Ко всем частям мозга были подключены детекторы спайков и вольтметры.

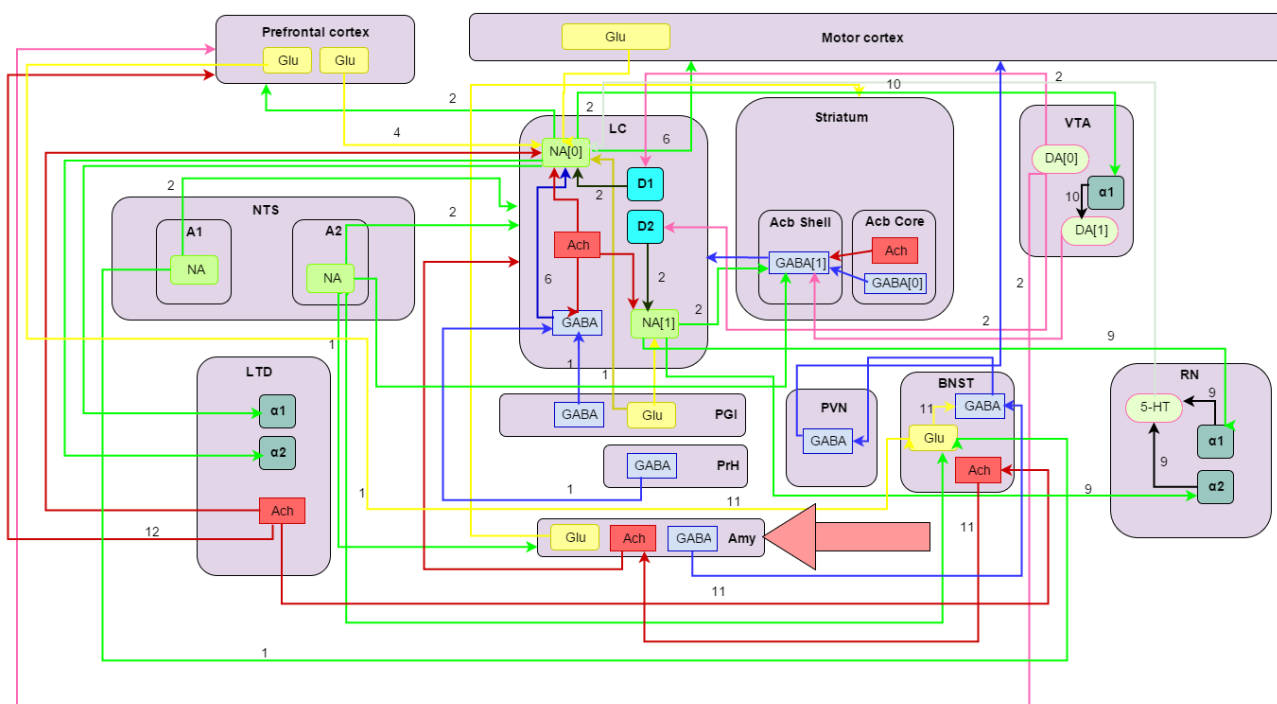


Рисунок 9. Схема распространения нейромедиатора норадреналин в мозге. Зелёными стрелками обозначен норадреналин, жёлтыми — глутамат, красными — ацетилхолин, синими — ГАМК, розовыми — дофамин. LC – голубоватое пятно, VTA - вентральная область покрышки, NTS – ядро одиночного пути, PGI - ядра paragigantocellular, PrH - perirhinal кора, LDT – латерально- спинное ядро покрышки, RN – ядра шва, BNST - слой ядра бороздки на терминале, Striatum – стриатум, Prefrontal cortex –префронтальная кора, Motor cortex –двигательная кора, Amy – миндалевидное тело, Thalamus – таламус, PVN – паравентрикулярное ядро.

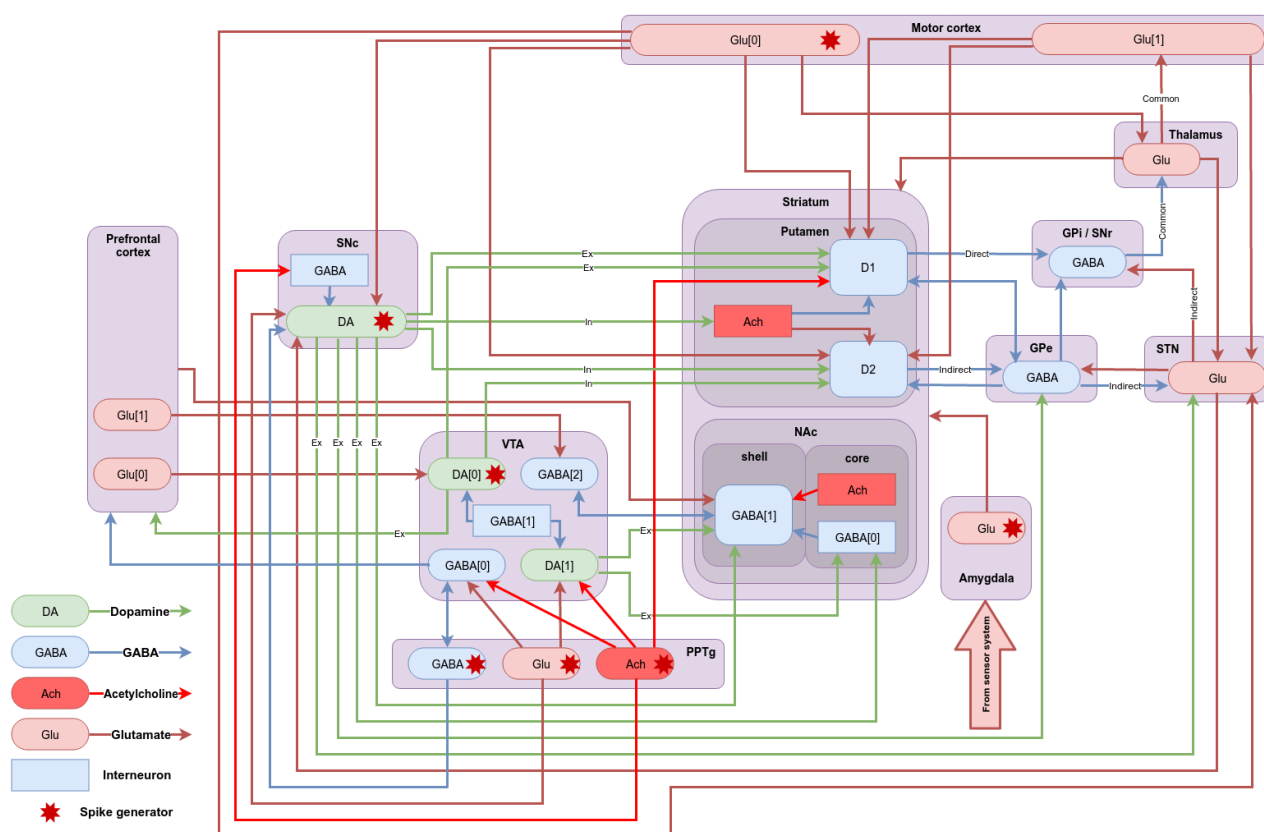


Рисунок 10. Схема распространения нейромедиатора дофамин в мозге. Motor cortex – моторная кора, Striatum – полосатое тело, GPe – продольный бледный шар, GPi – срединный бледный шар, STN – субталамическое ядро, Snc – черная субстанция pars compacta, Snr – черная субстанция pars reticulata, Amygdala – миндалевидное тело, Thalamus – таламус.

При объединении трёх систем возникли следующие вопросы:

- 1) Будут ли работающие по предложенной модели нейромедиаторы вмешиваться в работу друг друга на уровне отдельных нейронов;
- 2) Как поведёт себя участок модели мозга при чрезмерной стимуляции.
- 3) Каким образом балансировать сильное ингибирующее влияние серотонина на систему с возбуждающим действием дофамина, который в опасности быть полностью подавленным;
- 4) Какой компьютер выдержит нагрузку в 500 000 нейронов с многочисленными связями и как долго продлится одна подобная симуляция.

Для ответа на первый вопрос были проведены симуляции на простой сети из пяти нейронов: пресинаптический нейрон, постсинаптический нейрон, норадренергический нейрон, нейрон серотонина и нейрон норадреналина. К каждому нейрону были подключены детекторы спайков, вольтметры и генераторы спайков. Были проведены симуляции с включением одного из трёх нейромедиаторов при работающем другом нейромедиаторе, во всех комбинациях. Во всех случаях картина спайковой активности и динамики потенциала мембраны первого нейрона не изменялась, следовательно работа нейромедиаторов не саботирует друг друга на уровне отдельных нейромодулирующих нейронов. Пример показан на рис. 11: при включении на 600 миллисекунде серотонина, потенциал мембраны норадреналина продолжает находиться на уровне потенциала покоя, иногда происходит классический спайк. Концентрация норадреналина при этом подскакивает до максимума (рис. 12), поскольку событие «включение серотонина» для системы полностью неожиданное. На рис. 13 показана динамика работы пресинаптического и постсинаптического нейронов, нейромодулировать которые призваны норадреналин и серотонин (происходит активное спайкование).

На вопрос о чрезмерной стимуляции отдельного участка мозга ядро NEST версии 12 дало неожиданный результат: пробный запуск модели в 1000 нейронов дал нереалистичное изменение потенциала нейронов в миндалевидном теле (рис. NN). Потенциал закикливается вокруг -20 мВ, что не является потенциалом покоя и противоречит биологии, подобная ошибка стала возможна лишь

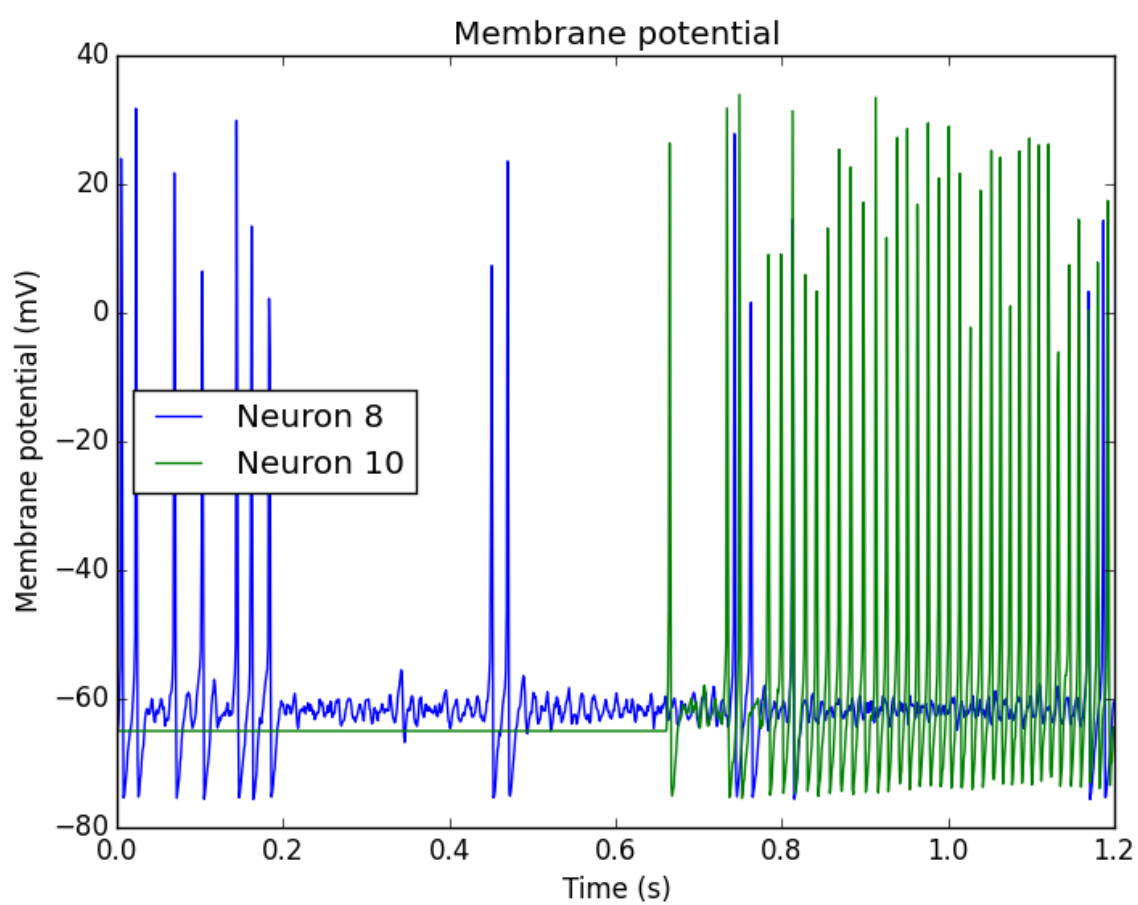


Рисунок 11. Изменение потенциала мембраны норадренергического нейрона (синий) и подключенного на 600 миллисекунде нейрона серотонина (зелёный).

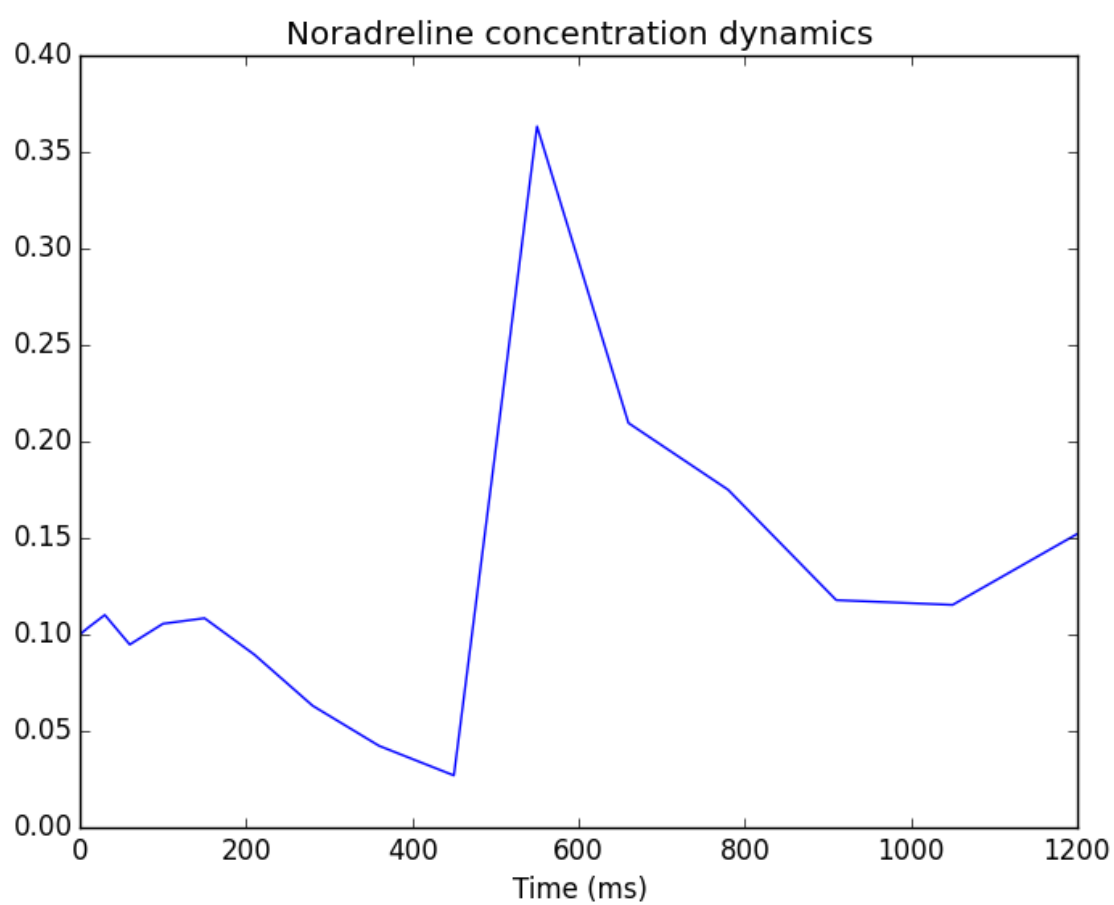


Рисунок 12. Изменение концентрации норадреналина при включении серотонина на 600 секунде.

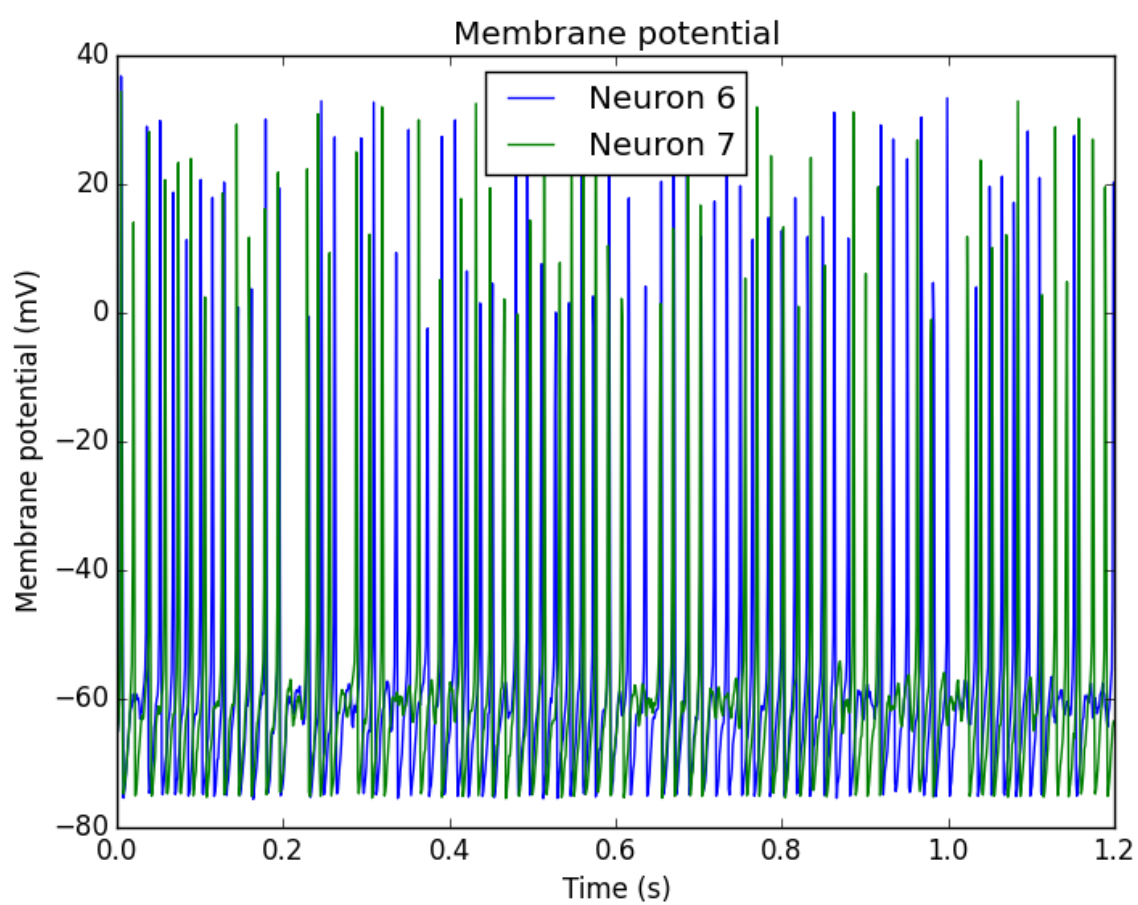


Рисунок 13. Изменение потенциала мембраны пресинаптического и постсинаптического нейронов при нейромодуляции.

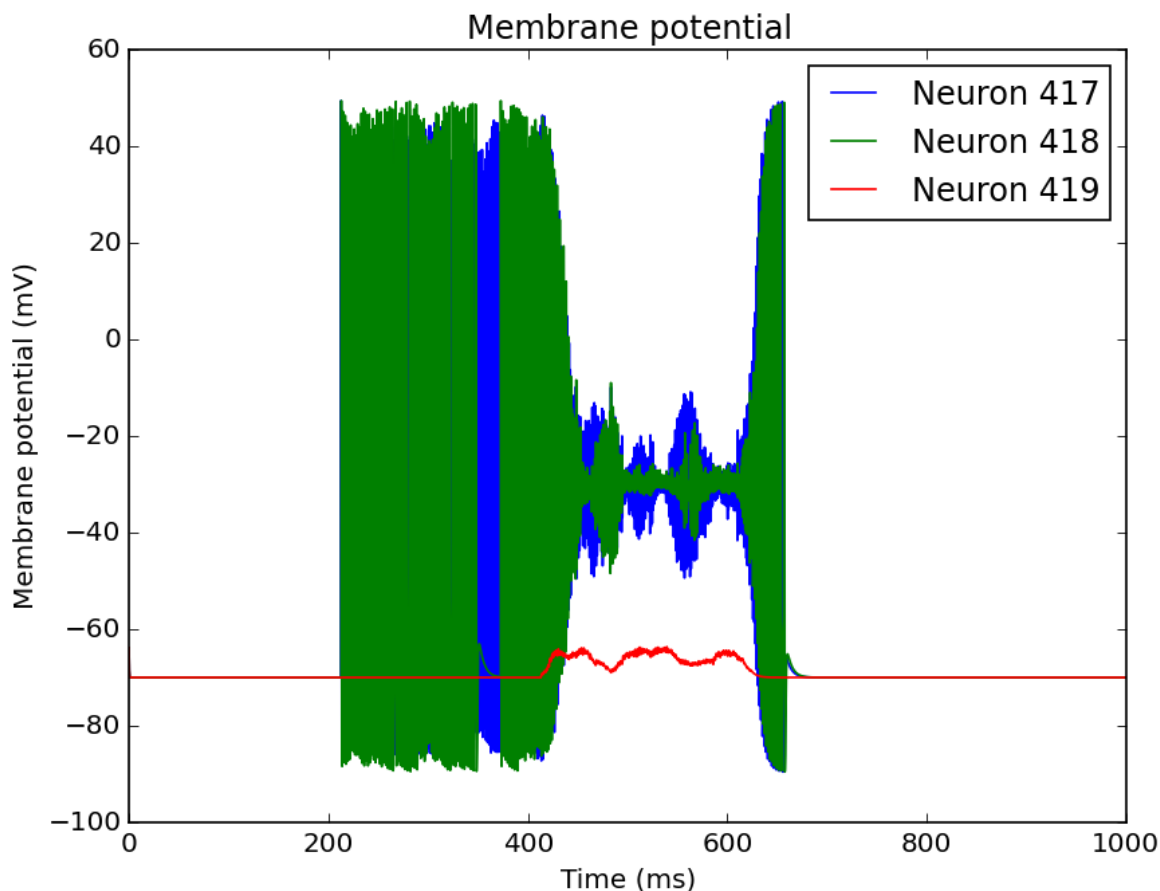


Рисунок 14. Устранённая мутация потенциала мембраны перестимулированного нейрона в нейронах миндалевидного тела (зелёный, синий).

из-за того, что в симуляции нейроны не сохраняют химические свойства, являются объектами с математическими характеристиками. Так как вопрос стоит на уровне версии NEST, то на данный момент он временно обойдён сокращением объёмов стимуляции у перестимулированных зон.

Вопрос о балансировании влияния серотонина на дофамин становится определяющим для реализации таких вычислительных эмоций, как радость и интерес. По модели Левхейма, эти эмоции возникают при высоких уровнях серотонина и дофамина, однако ингибирующее действие серотонина на дофамин почти исключает возможность их одновременной активности. Стало необходимо добавить новые центры дофамина

ГЛАВА 5. СИМУЛЯЦИЯ БАЗОВЫХ ЭМОЦИЙ В ТРЁХМЕРНОЙ МОДЕЛИ

В системе, полученной из трёх совмещённых систем нейромодуляторов на языке Python, с общим словарём областей мозга, выделенными для них нейронами в реалистичных пропорциях и настроенными связями, были проведены 8 симуляций на 500 000 нейронов с различными конфигурациями:

1. Эмоция «тоска/горе». Требуется высокого уровня норадреналина при низких уровнях дофамина и серотонина. Генераторы спайков подключены к ядрам солитарного тракта, ядрам *paragigantocellular*, латерально-спинному ядру покрышки и *perirhinal cortex* с 200 по 600 мс.

////////////////////////здесь будут картинки и объяснения картинок

2. Эмоция «презрение/отвращение». Требуется высокого уровня серотонина при низких уровнях дофамина и норадреналина. Генераторы спайков с 300 по 600 мс подключены к серотониновым очагам в префронтальной коре, моторной коре и ядрах шва (дорсальном и среднем).

////////////////////////здесь будут картинки и объяснения картинок

3. Эмоция «испуг/ужас». Формируется при высоком уровне дофамина и низких уровнях серотонина и норадреналина. Генераторы спайков подключены к моторной коре, чёрной субстанции *pars compacta*, миндалевидному телу и вентральной области покрышки с 200 по 600 мс.

////////////////////////здесь будут картинки и объяснения картинок

4. Эмоция «радость/удовольствие». Требуется одновременного существования спайковой активности в центрах дофамина и серотонина (чтобы серотонин не подавил дофамин), для чего были созданы центры — в —. Концентрация норадреналина должна оставаться низкой

5. Эмоция «злость/ярость». Возникает при высоких уровнях норадреналина и дофамина, а активность одного из них всегда влечёт активность другого. С 100 по 600 мс генерируются спайки на ядрах солитарного тракта, ядрах *paragigantocellular*, латерально-спинном ядре покрышки и *perirhinal cortex*. На 400 мс к ним добавляются генераторы на моторной коре, чёрной субстанции

pars compacta, миндалевидном теле и вентральной области покрышки, вместе они работают до 600 мс.

6. Эмоция «удивление». Высокие уровни серотонина и норадреналина, норадреналин борется за сохранение своего возбуждающего влияния, но не в полной мере сохраняет его под ингибирующим действием серотонина. К генераторам в префронтальной коре, моторной коре, дорсальном и среднем ядрах шва (работают с 300 по 600 мс) добавляются на 400 секунде генераторы на ядрах солитарного тракта, ядрах paragigantocellular, латерально-спинному ядру покрышки и perirrhinal cortex, вместе они работают до 600 мс.

7. Эмоция «стыд/унижение». Пониженные уровни всех трёх нейромедиаторов.

8. Эмоция «интерес/азарт». Высокие уровни всех трёх нейромедиаторов. Генераторы спайков в центрах дофамина (моторной коре, чёрной субстанции pars compacta, миндалевидном теле и вентральной области покрышки) работают начиная с 200 мс, генераторы в центрах серотонина (префронтальная кора, моторная кора, дорсальное и среднее ядра шва) подключаются на 300 мс, генераторы активности норадреналина (на ядрах солитарного тракта, ядрах paragigantocellular, латерально-спинном ядре покрышки и perirrhinal cortex) подключаются на 400 мс.

ЗАКЛЮЧЕНИЕ

Дипломная работа посвящена актуальной задаче реализации биологически реалистичных эмоциональных состояний на уровне вычислительных процессов, сопоставимых с мыслительными процессами в мозге млекопитающего. В рамках проведённых исследований выполнено:

- Реализована система распространения нейромодулятора норадреналин;
- Совмещены программные реализации систем норадреналина, серотонина и дофамина;
- Проведены эксперименты по исследованию воздействия трёх моноаминовых нейромодуляторов друг на друга, на связи между нейронами и на нейронную активность в разных областях мозга;
- Это воздействие воспроизведено на вычислительной нейронной сети в 500 000 нейронов (биологически реалистичное количество для мыши);
- Полученная система запущена в разных конфигурациях, согласно модели Хьюго Левхейма, итогом чего становятся зафиксированные и проанализированные 8 различных эмоций машины.

Результаты экспериментов показывают, что реалистичная симуляция функций биологических нейромодуляторов осуществима на вычислительных машинах, что становится новым шагом в области когнитивных технологий. Достигнута поставленная для дипломной работы цель: получена общедоступная, универсальная и эффективная модель вычислительных эмоций. Поскольку модель работает на уровне вычислительных процессов, её возможно перенести на аппаратное обеспечение, для объединения систем или модулей системы без необходимости перепрограммировать всю архитектуру. Реализованные в работе базовые эмоции имеют исключительно функциональные задачи: улучшения качества реагирования на внешние события, приспособляемости к внешним условиям, обучаемости и принятия решений.

Итогом исследования стало написание трёх статей по тематике работы:

- «NeuCogAr: how to make a machine feel emotions. Neuromodulating cognitive architecture for mammalian emotions simulation» в соавторстве с Хьюго Ле-

вхеймом, М. Талановым, J. Vallverdu, Б. Пинусом, Ф. Гафаровым, С. Дистефано, А. Хасьяновым, А. Тощевым, Е. Магидом, Р. Гайсиным и др. для журнала «Nature».

- «The Implementation of Noradrenaline in the NeuCogAr Cognitive Architecture» в соавторстве с М. Талановым, Б. Пинусом, J. Vallverdu и др. для участия в IX международной конференции по продвинутым когнитивным технологиям и приложениям COGNITIVE-2017.
- «Extending biomimetic cognitive architecture NeuCogAr with noradrenaline model» в соавторстве с М. Талановым, Б. Пинусом, J. Vallverdu, С. Дистефано, Ф. Гафаровым и А. Леухиным для журнала «Journal of Healthcare Engineering».