

ОБЗОРНЫЕ И ПРОБЛЕМНЫЕ СТАТЬИ

DOI: 10.7868/S0869813918100015

**НЕЙРОНАЛЬНАЯ МОДЕЛЬ АДАПТИВНОГО
ПОИСКОВОГО ПОВЕДЕНИЯ**

© С. В. Альбертин

Институт физиологии им. И. П. Павлова РАН, Санкт-Петербург, Россия
E-mail: salber1@rambler.ru

В статье рассматриваются возможности разработки и использования различных нейронных моделей целенаправленного поведения животных при создании систем искусственного интеллекта. Представлена нейронная модель адаптивного поведения, основанная на использовании радиального лабиринта с дифференцированным пищевым подкреплением. Обсуждаются ее преимущества по сравнению с нейронными моделями целенаправленного избегательного поведения животных в водных лабиринтах.

Ключевые слова: адаптивное поведение, нейронное моделирование, память, биоподобные роботы, искусственный интеллект.

Рос. физиол. журн. им. И. М. Сеченова. Т. 104. № 10. С. 1121—1135. 2018

S. V. Albertin. NEURAL MODEL OF ADAPTIVE SEEKING BEHAVIOR. Pavlov Institute of Physiology of the RAS, St. Petersburg, Russia; e-mail: salber1@rambler.ru.

The article discusses the advantages of developing and using of various neuronal models of animal goal-directed behavior to create artificial intelligence systems. A new neuronal model of adaptive behavior is presented, based on the use of a radial maze with differentiated food reinforcement. Its advantages in comparison with neuronal models of purposeful avoidance behavior of animals in water mazes are discussed.

Key words: adaptive behavior, neuronal modeling, memory, biosimilar robots, artificial intelligence.

RUSSIAN JOURNAL OF PHYSIOLOGY. V. 104. N 10. P. 1121—1135. 2018

В последние годы наблюдается резкий рост исследований, направленных на моделирование и конструирование искусственно созданных автономных устройств — биоподобных электромеханических роботов, способных к обучению, ориентации в пространстве, целеполаганию и целенаправленному поведению [19, 30, 39]. При создании биоподобных автономных устройств исследователи обычно используют различные нейронные модели целенаправленного поведения, выполненные на животных, это позволяет исследовать нейрофизиологические механизмы, стратегию и тактику поведения испытуемых при изменении условий реагирования, сетевую архитектуру и алгоритм

осуществления отдельных поведенческих актов и на их основе конструировать биоподобных и антропоморфных роботов, способных копировать адаптивное поведение различных видов животных [11, 12, 22, 31, 32, 39].

Одной из таких моделей является нейрональная модель поведения крыс в водном бассейне, которая основана на поиске животными, помещаемыми в бассейн, местонахождения погруженной в воду безопасной платформы, определяемой с помощью экстралабиринтных навигационных знаков — лабиринт Морриса [36]. На решение задачи всем тестируемым животным отводится одинаковое время. При появлении во время поиска платформы опасности гибели (утопления) животного выполнение задачи прерывают и крысу возвращают на одну из стартовых позиций. Разновидностью лабиринта Морриса является «сухая» версия водного лабиринта, известная как лабиринт Барнеса [41], позволяющая животным избегать выбора заполненных водой локусов, расположенных по окружности арены, при поиске единственного пустого локуса, служащего животному укрытием от подаваемых экспериментатором аверсивных внешних воздействий (ярких световых импульсов). Ориентация животных в этом случае также осуществлялась с помощью внешних навигационных знаков, расположенных вокруг экспериментальной арены.

Используя результаты анализа нейрональной активности гиппокампа и энторинальной коры, регистрируемой при осуществлении поискового поведения крыс в водном лабиринте Морриса, а также архитектуры и алгоритма осуществления отдельных поведенческих актов J. L. Krichmar и соавт. [31, 32] создали сетевую нейроморфную модель целенаправленного поискового поведения. На основе этой модели авторы сконструировали электромеханического биоподобного робота «Darwin X», оснащенного сенсорными датчиками, который был бы способен находить в условиях так называемой сухой версии лабиринта скрытую (hidden) платформу с помощью внешних зрительных ориентиров, копируя целенаправленное поведение подопытных животных в водных лабиринтах. Следует отметить, что условия тестирования сконструированного робота существенно отличались от обучения и тестирования подопытных животных в лабиринте Морриса. Бассейн округлой формы, заполненный водой, при тестировании робота заменялся сухой прямоугольной площадкой (4.88/4.27 м) с ориентирами в виде полос разного цвета, расположенными по сторонам рабочей зоны площадки [31, 32]. На площадке размещалась так называемая «скрытая» платформа, выполненная в виде круга диаметром 61 см покрытого отражающим материалом, который мог быть обнаружен роботом с помощью встроенных инфракрасных датчиков, сканирующих при движении робота исследуемую поверхность площадки. В случае неудачи обнаружения платформы в течение фиксированного времени (17 мин) поиск прерывали и робот возвращали на одну из 4 стартовых позиций, расположенных в различных частях рабочей площадки. Успешный поиск роботом «скрытой» платформы в этих экспериментах оценивался исследователями по количеству правильных решений, а также по наличию сходных рисунков маршрутов поиска, выявленных у подопытных животных и тестируемого робота. Результаты тестирования проведенного создателями анимата «Darwin X» показали возможность правильного решения задачи роботом после 20 последовательно проведенных проб обучения.

Нейроморфная модель целенаправленного поведения в водном лабиринте Морриса, использованная J. L. Krichmar и соавт. [31, 32], показала принципиальную возможность создания биоробота, способного к правильному решению поисковой задачи при использовании так называемой сухой версии лабиринта. Вместе с тем следует отметить, что исходная поведенческая мо-

дель поискового поведения в водном лабиринте, используемая на животных при определении сетевой архитектуры и алгоритма поведенческих актов, необходимых для создания нейроморфной модели целенаправленного поведения, имеет следующие недостатки, которые ограничивают ее применение при конструировании и дальнейшем совершенствовании биоподобных автономных устройств.

1. Условия обучения и тестирования животного (water Morris maze) и созданного на его основе биоробота (dry-version space task) имеют значительные различия, которые затрудняют корректное сравнение регистрируемых параметров поведения и взаимную экстраполяцию полученных результатов исследований.

2. При тестировании крыс с использованием водного лабиринта Морриса или его разновидности — лабиринта Барнеса у животных наблюдаются выраженные индивидуальные, видовые (крысы, мыши) и генотипические различия регистрируемых нейрофизиологических и поведенческих показателей [41]. При этом многие исследователи отмечали у подопытных животных развитие выраженных стрессорных реакций [13, 26, 27], что может изменять нормальное восприятие предъявляемых навигационных сигналов и как следствие влиять на выбор и осуществление адекватного поведения.

3. Применение водных лабиринтов с аверсивными видами воздействия может вызывать патологические изменения регистрируемой нейрональной активности гиппокампа и связанных с ним структур мозга. Это может существенно искажать нормальную картину сетевого нейронального реагирования, алгоритм и маршруты выполняемых поисковых двигательных реакций животных, что затрудняет создание биологически полноценной нейроморфной модели поискового поведения подопытного животного, необходимой для конструирования биоподобных роботов.

4. Использование водных лабиринтов позволяет моделировать лишь однозначное, строго детерминированное целенаправленное поведение, направленное на избегание аверсивного воздействия, но исключает возможность усложнять условия проведения эксперимента, требующие от подопытных животных адаптивно применять разные тактики мотивированного поведения при решении поисковой задачи, которые возможны в условиях подкрепляемого пищевого поведения с выбором одного из различных видов предлагаемых вознаграждений.

Ниже нами представлены результаты исследований, проведенные с использованием альтернативной нейрональной модели адаптивного поискового поведения животных, которая разработана в Национальном центре научных исследований Франции (CNRS) и Институте физиологии им. И. П. Павлова РАН при исследовании поискового поведения крыс в радиальном лабиринте с дифференцированным пищевым подкреплением [14—16, 44].

МЕТОДИКА ИССЛЕДОВАНИЯ

В соответствии с разработанной моделью адаптивного поискового поведения подопытные животные (крысы линии Long Evans) были обучены заходить в один из отсеков 4-рукавного радиального лабиринта (рис. 1), каждый из которых оснащен кормушкой для подачи в жидком виде подкрепления (вода, молоко) и электрической лампочкой, включение которой осуществлялось при выходе крысы на центральную площадку лабиринта и являлось пусковым условным сигналом для захода животного в подкрепляемый рукав

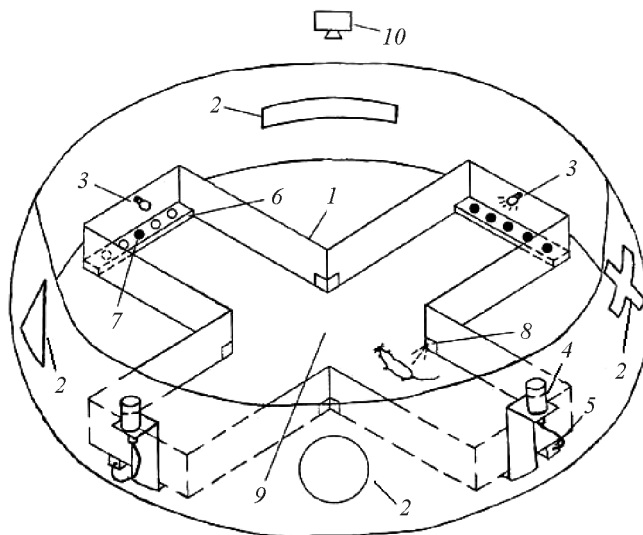


Рис. 1. Схема устройства лабиринта. 1 — радиальный 4-рукавный лабиринт, 2 — постоянно действующие зрительные ориентиры (геометрические фигуры), 3 — внутрилабиринтные целевые стимулы (лампочки), 4 — емкость с трубопроводом для хранения и подачи подкрепления, 5 — средства для подачи подкрепления (соленоид), 6 — кормушка с пятью питьевыми локусами, 7 — локусы кормушки с подкреплением, 8 — средства включения лампочки и одновременной подачи вознаграждения в подкрепляемом рукаве лабиринта (ИФК-датчики), 9 — центральная площадка лабиринта, 10 — видеокамера.

(pavlovian stimulus/response). Выбор животным правильного (освещенного) рукава сопровождался подачей дифференцированного подкрепления, при этом в трех рукавах крысы получали по одной капле, а в одном рукаве — пять капель подкрепления, которые подавались в дискретном режиме в центральной локус кормушки (последовательно по одной капле, интервал между каплями 1—5 с). Подача каждой капли подкрепления сопровождалась щелчками соленоида, воспринимаемыми подопытным животным. В отдельной серии экспериментов максимальное подкрепление подавалось одновременно по одной капле в пяти раздельных локусах кормушки, хорошо видимых животным при подходе к ней (рис. 1). Включения внутрилабиринтных зрительных сигналов (лампочек) осуществлялись в случайном порядке. Ошибочные ответы (заходы в неосвещенные отсеки) и попытки повторного подхода подопытных животных к одной и той же поилке без выхода на центральную площадку лабиринта не подкреплялись.

Способность подопытных крыс к запоминанию в пространстве местоположения отсека с наибольшим (5 капель) подкреплением тестировалась с помощью серии пробных тестов, *probe trials* — одновременным включением зрительных условных сигналов (лампочек) во всех 4 рукавах лабиринта, когда животным предоставляли возможность получить подкрепление в любом рукаве лабиринта (рис. 1, 2). Одновременное включение всех лампочек служило для животных сигналом для изменения стратегии поведения (Test 1 — *set shifting strategy*), заключающейся в переходе от реагирования на включение одного из внутрилабиринтных раздражителей — *simple stimulus/response*, так называемое *model free behavior* [21], к поиску предпочитаемого подкрепления в лабиринте с помощью конфигурации (модели) навигационных знаков —

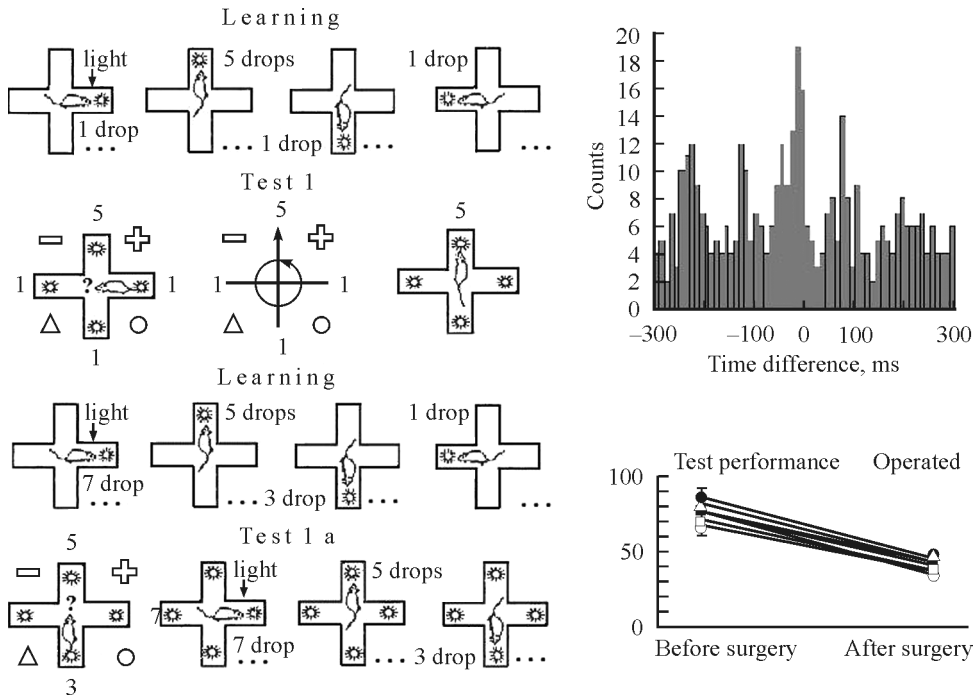


Рис. 2. Выбор подопытными животными предпочитаемого подкрепления в лабиринте с дифференцированным подкреплением при одновременном включении локальных зрительных сигналов во всех рукавах радиального лабиринта — Tests 1, 1a [14, 15].

Цифры на рисунке — количество капель подкрепления в рукавах лабиринта. Кружки со стрелками — направление обзорного движения животного при определении локализации предпочитаемого подкрепления в лабиринте с помощью экстралабиринтных навигационных сигналов.

Справа сверху — кросскоррелограмма синхронной активности в тета-диапазоне нейронной активности прилежащего ядра и гиппокампа, регистрируемой при выборе предпочитаемого (наибольшего) подкрепления. Справа внизу — уровень правильных ответов при выборе наибольшего подкрепления до и после повреждения прилежащего ядра.

хорошо видимых геометрических фигур, которые нанесены на внутреннюю поверхность ширмы, установленной вокруг лабиринта — Test 1, *model based behavior* [21].

Условнорефлекторные переделки (Tests 1.2, 1.3, 1.4 — *reversal learning*), связанные с изменением местоположения максимального подкрепления в рукавах лабиринта, проводились ежедневно утром после ночного отдыха животного.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ИХ ОБСУЖДЕНИЕ

Проведенные нами исследования показали, что интактные животные в 80—90 % тестируемых пробах (Tests 1, 1.2, 1.3, 1.4) при одновременном включении локальных зрительных сигналов выбирают рукав лабиринта с наибольшим подкреплением [14, 15]. При этом показано, что выбор предпочитаемого подкрепления у подопытных животных сопровождался синхронизацией нейрональной активности прилежащего ядра (n. accumbens) и гиппокампа в тета-диапазоне [16, 44]. Электролитическое повреждение прилежащего ядра

ра [15] приводило к нарушению способности животных к нахождению местоположения наибольшего подкрепления в лабиринте по навигационным ориентирам (Test 1), а также в последующие дни тестирования (Tests 1.2, 1.3, 1.4) при изменении местоположения максимального подкрепления в рукавах лабиринта (рис. 2). Сходные результаты нарушения поведения наблюдались у интактных подопытных животных, тестируемых после инъекции агониста амфетамина [6]. В отдельной серии экспериментов проводили обучение животных с дифференцированным (1, 3, 5, 7 капель) подкреплением, подаваемым во всех 4 рукавах радиального лабиринта — Test 1A (рис. 2). Обеспечение свободного доступа в тестовой пробе (test probe) к кормушкам во всех рукавах лабиринта показало, что животные последовательно выбирали рукава лабиринта с 7, 5, 3 и 1 каплями, сохраняя приоритет выбора наибольшего подкрепления, что свидетельствовало о возможности формирования у животного карт памяти о локализации всех предъявляемых подкреплений (рис. 2). При этом у этих животных, так же как и при выполнении предыдущей задачи с дифференцированным подкреплением (Test 1), сохранялось наличие синхронной активности гиппокампа и прилежащего ядра при выборе каждого подкрепляемого рукава [42].

Используя вышеописанную нейрональную модель адаптивного поискового поведения крыс в радиальном лабиринте с дифференцированным подкреплением, наши коллеги, работавшие в Национальном центре научных исследований Франции [28, 35], а также исследователи Шеффилдского университета Великобритании [29] сконструировали электромеханических аппаратно-программных роботов, предназначенных воспроизводить алгоритм вышеописанного поведения: правильно реагировать на изолированные включения внутрилабиринтных раздражителей (pavlovian stimulus/response test) и переключать внимание на внешние навигационные знаки при изменении стратегии поведения (Test 1 — set shifting strategy), наблюдаемого у реальных крыс в 4-рукавном лабиринте [14–16].

Биороботы представляли собой стилизованные в виде крыс электромеханические устройства, установленные на подвижной платформе, оснащенные двигателем с тахометром для регистрации скорости движения робота, контроллерами ходовых двигателей и исполнительных механизмов, бинокулярными зрительными и бинауральными слуховыми сенсорами, тактильными тензодатчиками в виде вибрисс и сервоприводами, позволяющими изменять положение головы и туловища биоподобного робота (рис. 3). При этом изменения положения и повороты робота регистрировались с помощью роторного сенсора, копирующего функции вестибулярного аппарата животного. Вознаграждением при правильном выборе рукава лабиринта служила энергетическая подзарядка аккумулятора робота, которая позволяла роботу отмечать, когда и какое количество подкрепления получено.

Сконструированные нашими коллегами биоподобные роботы в целом могли правильно осуществлять выбор рукава лабиринта, ориентируясь на внутрилабиринтные сенсорные сигналы, переключать внимание на внешние навигационные знаки при изменении стратегии поведения, а также находить наибольшее подкрепление (Tests 1.2, 1.3, 1.4 — *reversal learning*) при изменении его местоположения в рукавах лабиринта.

Вместе с тем исследователи [28] отмечали, что скорость выполнения задачи у биороботов по сравнению с подопытными крысами была намного ниже, чем у реальных животных, несмотря на то что вибриссы-тензодатчики помогали им успешно обходить встречающиеся на пути препятствия (стенки и углы лабиринта). Это объясняется тем, что в отличие от биороботов при перво-

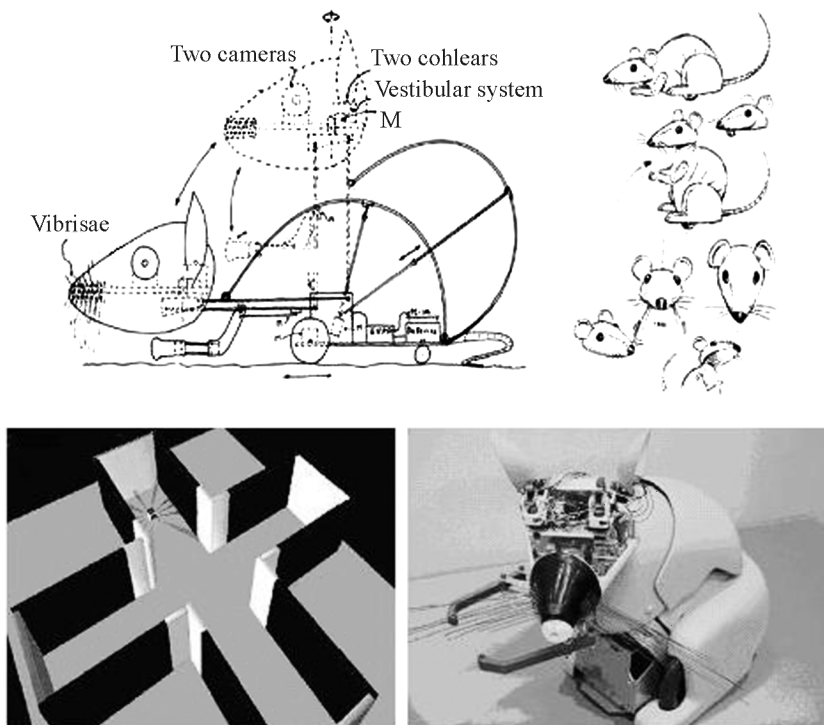


Рис. 3. Схематическое устройство и общий вид аппаратно-программного биоробота Psikharrax [28, 35], выполненного на основе предложенной нейрональной модели поискового поведения крыс в радиальном лабиринте с дифференцированным подкреплением [14, 15]. Справа вверху — возможные позиции головы и туловища биоробота при осуществлении поисковых двигательных реакций в лабиринте.

начальном обучении реальных животных (pavlovian stimulus/response) фактически задействовано два вида памяти: подкрепляемое ассоциативное S-R обусловливание (сигнал—ответ) и параллельно протекающее латентное обучение, основанное на формировании пространственной памяти о локализации наибольшего подкрепления, которое обеспечивало подопытным крысам быстрое и практически безошибочное решение задачи при тестировании последующего поведенческого акта (Test 1). У сконструированных биороботов вышеуказанное параллельное обучение вызывало очевидные затруднения.

Следует отметить, что вышеописанные биороботы [28, 35] были сконструированы для воспроизведения относительно простой формы павловского ассоциативного обучения (подача внутрилабиринтного сигнала — побежка в сигнализируемый рукав радиального лабиринта), усложненной последующей сменой стратегии поведения, связанной с использованием пространственной памяти при выборе подкрепления с помощью экстралабиринтных навигационных ориентиров — Test 1 [15]. Однако они не были приспособлены адаптивно менять свое поведение при изменении других условий реагирования в радиальном лабиринте (требующих использования различных видов памяти), которые были подробно исследованы в наших следующих экспериментах.

Дальнейшие исследования, проведенные на крысах в радиальном лабиринте с асимметричным подкреплением [4–9], связанные с изменением условий реагирования в лабиринте (Tests 2, 3), показали, что наряду с ассоциатив-

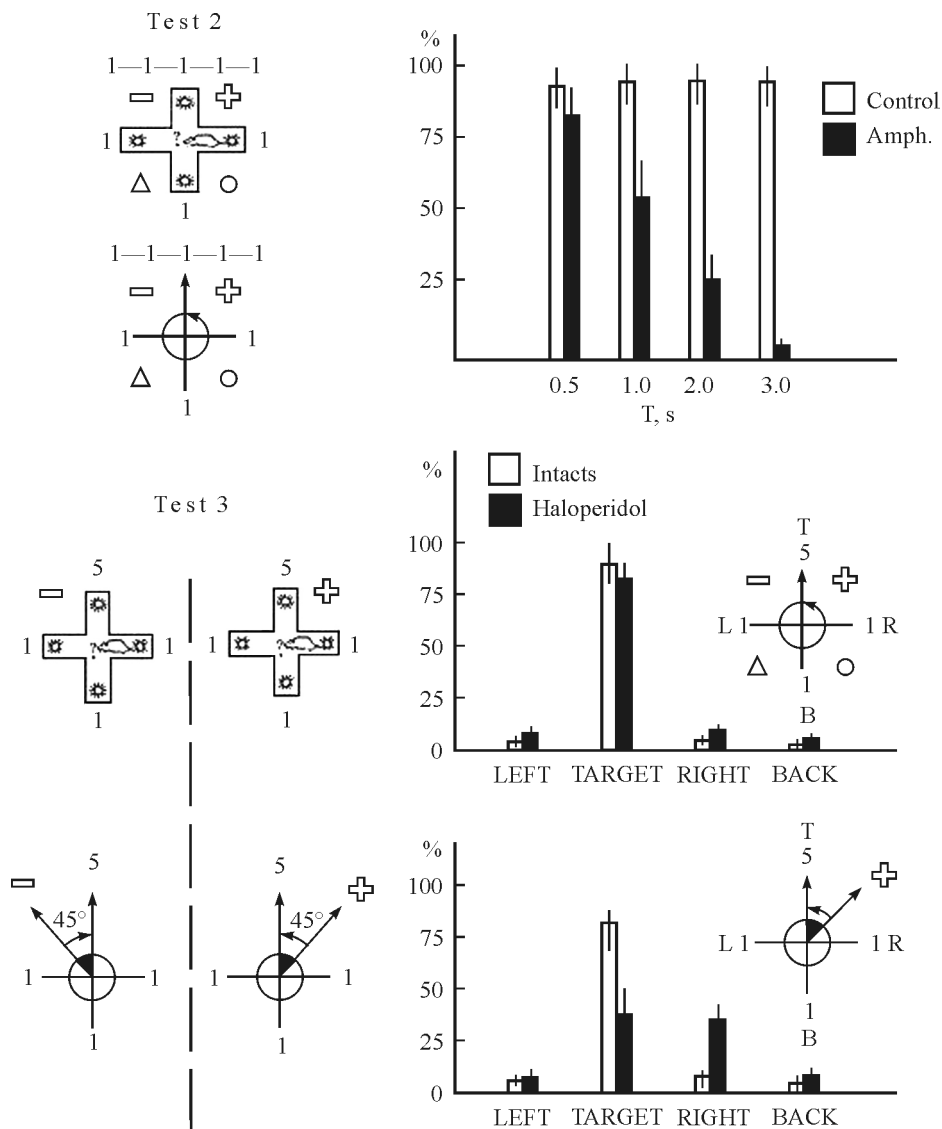


Рис. 4. Тестирование подопытных животных до и после инъекции ДАагониста амфетамина при увеличении отсрочки между дискретно подаваемыми порциями наибольшего (5 капель) подкрепления — Test 2 [4, 5]. Тестирование животных до и после инъекции ДАантагониста галоперидола при наличии лишь одного из двух смежных экстралабиринтных зрительных навигационных ориентиров — Test 3 [9].

Цифры на рисунке — количество капель подкрепления в рукавах лабиринта. Кружки со стрелками — направление движения животного при определении локализации предпочитаемого подкрепления в лабиринте с помощью экстралабиринтных навигационных сигналов.

Справа сверху — число (%) ответов с выбором наибольшего подкрепления при увеличении времени между дискретно предъявляемыми порциями подкрепления. Справа внизу — количество (%) выборов различных рукавов лабиринта при полной и фрагментарной экспозиции навигационных ориентиров (target — целевой рукав с наибольшим подкреплением, left — левый, right — правый, back — противоположный целевому рукава лабиринта).

ным обучением (стимул/ответ) и краткосрочной пространственной памятью (Test 1) животные успешно используют возможности других адаптивных тактик, основанных на применении различных видов памяти (рис. 4, 5).

Показано, что при увеличении отсрочки между дискретно подаваемыми порциями наибольшего (5 капель) подкрепления до 3—5 с (рис. 4, Test 2) интактные животные, используя возможности краткосрочной памяти — Short-term Memory (STM), были способны к ожиданию следующей отсроченной порции подкрепления, однако эта способность нарушалась при стимуляции ДА-реактивной системы вентрального стриатума [4, 5].

В другом исследовании (рис. 4, Test 3) тестирование проводилось без отсрочки подкрепления, но при фрагментарной экспозиции экстралабиринтных навигационных знаков [9]. Три геометрические фигуры убирала, оставляя один из двух парных сигналов, которые ранее однозначно определяли локализацию предпочитаемого подкрепления. При фрагментарной экспозиции навигационных знаков животному предъявляли один из смежных экстралабиринтных сигналов, которые чередовали в тестируемых пробах в случайном порядке, чтобы исключить возможность обучения животных выполнению двигательной реакции одностороннего поворота головы и туловища, которые бы позволили безошибочно выбирать подкрепляемый рукав путем использования эгоцентрической тактики поведения, компенсируя тем самым возможные нарушения памяти животных (рис. 4). Интактные животные сохраняли способность к инвариантному восприятию прежней конфигурации ориентиров, необходимой при определении наибольшего подкрепления с помощью активного использования эпизодической памяти — Episodic Short-term Memory (ESM), которая, однако, нарушалась при дисбалансе ДАергической системы вентрального стриатума [9]. Показано, что основой выявленных нарушений является дефицит ассоциативной памяти между оставшимся и удаленным навигационными знаками. Предполагается, что в процессах активации ассоциативной связи между значимыми смежными экстралабиринтными сигналами, которые определяют локализацию рукава лабиринта с наибольшим подкреплением, и вытормаживание конкурирующих ассоциативных связей, не обеспечивающих получение предпочитаемого подкрепления наряду с тесными гиппокамально-стриатными взаимоотношениями [44], важную роль может играть динамика когерентной нейрональной активности гиппокампа и префронтальной коры [18, 25]. При этом в электрофизиологических исследованиях показано, что дофамин оказывает модулирующие влияния на синаптическую активность, процессы долговременной потенциации и когерентную нейрональную активность указанных структур мозга, необходимую для формирования эпизодической памяти [20, 43].

Наличие выявленной у животных эпизодической памяти дает им возможность адаптивно реагировать при изменении условий реагирования в лабиринте, так как позволяет инвариантно воспринимать неполные или зашумленные фрагменты окружающей обстановки и принимать решения в обстановке, сличая ее со своим прошлым опытом. Следует отметить, что эпизодическая память является основой для формирования автобиографической памяти человека, дающей возможность безошибочно идентифицировать не только объекты, но и различные субъекты, включая выражение лица и мимики человека, являющиеся необходимым условием для социальной коммуникации, а также позволяет восстанавливать в памяти контекстуальные детали различных сцен.

Особый интерес вызывают результаты наших исследований, связанные с экстренным переобучением подопытных животных поиску наибольшего под-

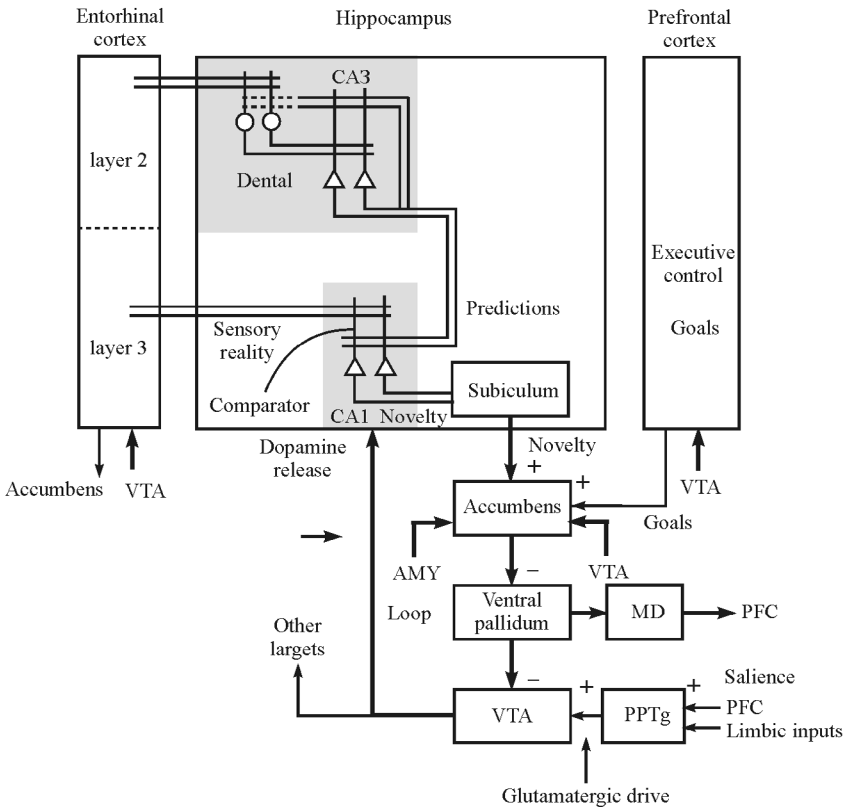
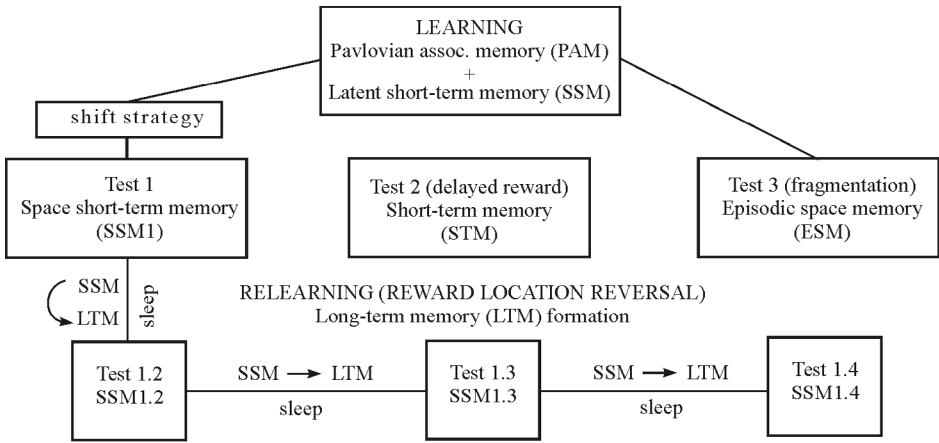


Рис. 5. Взаимодействие различных видов памяти при ориентации крыс в лабиринте (верхний рисунок). Модифицированная схема нейрональной организации структур головного мозга [33], участвующих в формировании адаптивного поведения при изменении условий реагирования в радиальном лабиринте (нижний рисунок).

VTA — ДАергические проекции вентральной тегментальной области среднего мозга; PFC — префронтальная кора; AMY — амигдала; CA1, CA3 — поля гиппокампа; PPTg — ножково-мостовое ядро тегмента; MD — медиодорсальное ядро таламуса. Активирующие и тормозные влияния соответственно обозначены знаками «плюс» и «минус».

крепления при изменении его локализации в радиальном лабиринте, определяемого с помощью новой конфигурации (моделью) навигационных ориентиров [7]. Известно, что при переобучении животных морфофункциональной основой запечатленных пространственно-организованных моделей окружающей его среды является формирование констелляций (пулов) нейронов со специфическими паттернами активности, которые могут сохраняться в течение длительного времени или изменяться (дополняться, замещаться, стираться, разрушаться) в зависимости от получения новых знаний об изменении окружающей среды [4].

Попытки в наших исследованиях провести повторные, выполненные без перерыва в день опыта, сеансы переобучения животных нахождение нового места предпочитаемого подкрепления в радиальном лабиринте приводили у интактных подопытных животных к нарушению способности использовать навигационную ориентацию (model-based behavior) как более адаптивную тактику для максимизации подкрепления при сохранении менее продуктивного условнорефлекторного реагирования на подачу внутрилабиринтных сигналов [7]. Фармакологические (введение испытуемым препарата, индуцирующего сон) и процедурные (предоставление в перерыве между сеансами с переобучением кратковременного отдыха со сном) воздействия свидетельствовали о функциональной (обратимой) природе нарушений, выявленных у подопытных животных [7]. Минимизация этих нарушений у животных после естественного или вызванного введением нейропептида сна подтверждает ранее высказанное предположение о необходимости нейрональных перестроек мозга во время сна для консолидации следов памяти (их переводу в долгосрочную форму памяти) и восприятия нервными клетками новой информации об изменениях в окружающей среде [23, 34, 38, 40].

Результаты проведенных исследований показали, что в организации исследованных форм адаптивного поведения в радиальном лабиринте с дифференцированным пищевым подкреплением животные наряду с простой формой ассоциативного обучения Pavlovian Associative Memory (PAM) активно используют различные виды памяти: пространственную краткосрочную память — Space Short-term Memory (SSM), эпизодическую память при фрагментарной экспозиции сигналов — Episodic Short-term Memory (ESM), краткосрочную память — Short-term Memory (STM) на задержанное подкрепление и долговременную память — Long-term Memory (LTM), на основе которых принимаются решения о выборе адаптивной стратегии поведения и соответствующей реакции (рис. 5). При этом выявлен алгоритм и сложная архитектура поведения крыс с различными уровнями принятия решения о выборе адаптивной стратегии при ориентации в лабиринте, в организации которых ведущую роль играет вентральный стриатум (*n. accumbens*), получающий прямые моносинаптические проекции от ДАергических нейронов вентральной тегментальной области среднего мозга (VTA), гиппокампа (Subiculum), энторинальной коры (EC), амигдалы (AMY), префронтальной коры (PFC) и посылающий эффекторные проекции к вентральному паллидumu, который в свою очередь посылает прямые проекции к нейронам VTA и медиодорсальному ядру таламуса (MD), связанному с префронтальной корой (рис. 5). Отметим, что указанные структуры мозга связаны с обеспечением функциональной активности петли, ведущая роль в которой принадлежит гиппокампу и ДАергическим нейронам среднего мозга (VTA). Вход в зубчатую извилину гиппокампа идет из слоя 2 энторинальной коры. Область зубчатой извилины и поля CA3 гиппокампа хранят и воспроизводят тестируемые последовательности и обеспечивают вход к полю CA1 гиппокампа, который сравнивает

эти предположения с сенсорной информацией, поступающей из 3 слоя энторинальной коры. Сигнал новизны из компаратора CA1 приходит к субикулюму, затем к базальным структурам переднего мозга, а оттуда к ДА-клеткам вентрального тегмента среднего мозга (VTA). Петля завершается восходящими ДА-проеекциями, которые иннервируют гиппокамп. Активирующие лимбические влияния поступают в VTA из ножково-мостового ядра тегмента (PPTg).

Проведенные нами эксперименты выявили особую роль ДАергической системы при реализации адаптивного поискового поведения в радиальном лабиринте с дифференцированным подкреплением. Дисфункция этой системы, вызванная инъекцией агонистов и антагонистов дофамина, приводила к нарушению выполнения различных поведенческих тестов. Выявленные эффекты влияния дофамина обусловлены не только наличием афферентных влияний ДАергической системы среднего мозга на гиппокамп, прилежащее ядро, префронтальную и энторинальную кору — основные структуры головного мозга, когерентная активность которых необходима при формировании различных видов памяти и реализации пространственной ориентации на тестируемые сигналы (рис. 5), но также высокой мотивацией и эмоциональной напряженностью последовательно выполняемых поведенческих актов, которая в наших опытах усиливалась по мере последовательного приближения (*successive approaching behavior*) животного к получению предпочитаемого подкрепления [4, 6]. При этом показано, что аттрактивность (*salience*) наибольшего (5 капель) подкрепления увеличивалась при переходе от дискретной по времени подачи отдельных порций подкрепления к одновременному предъявлению 5 капель (рис. 1), хорошо видимых животным при подходе к кормушке [8]. Важная роль ДАергической системы в реализации данной формы поведения (*approaching behavior*) ранее отмечалась исследователями также при реализации последовательно выполняемых инструментальных реакций, необходимых для получения подкрепления [4, 24].

Следует отметить, что биоподобные роботы [28, 29], выполненные на основе описанной нейрональной модели с дифференцированным пищевым подкреплением, могли максимизировать вознаграждение (подзарядка аккумулятора), которое можно было получать регулярно в течение длительного периода времени при правильном реагировании в лабиринте. При этом робот на основе показаний встроенного счетчика имел собственную оценку количества «заработанной» энергии аккумулятора, позволяющую роботу продлить свою активность. Таким образом, робот, являясь автономным устройством, имел потребность определять уровень заряда аккумулятора и при недостатке энергии изменять тактику поведения для получения дополнительного вознаграждения в виде подзарядки аккумулятора. Результаты ранее проведенных нейрональных исследований предполагают, что функцию «критика», оценивающего необходимость и возможность получения подкрепления животным, в описанной нами модели подкрепляемого обучения выполняет прилежащее ядро (*n. accumbens*), а функцию исполнителя продуктивной тактики поведения так называемого «актера» — дорсальный стриатум [28, 44].

Возникает вопрос, возможно ли сконструировать автономных биороботов, способных оперативно реагировать на все вышеописанные изменения условий реагирования (Тесты 1, 1а, 1.2, 1.3, 1.4, 2, 3), решаемые подопытными животными в лабиринте с дифференцированным подкреплением? Есть основания полагать, что дальнейшая аппаратно-программная модернизация биороботов путем их оснащения современными модульными электронными запоминающими устройствами, выполненная с учетом выявленной сложной

многоуровневой архитектуры и алгоритма поведения тестируемых крыс в лабиринте с дифференцированным подкреплением, позволит будущим биороботам сравняться в скорости реагирования с реальными животными, а в отдельных случаях (Тесты 1.2, 1.3, 1.4 — Reversal learning) превосходить скорость переобучения подопытных крыс, так как биороботы в отличие от реальных животных не требуют сна для нейронального перепрограммирования. Следует отметить, что многоуровневая архитектура поведения с использованием различных уровней принятия решения о выборе адаптивной стратегии поведения в настоящее время начинает активно использоваться в роботостроении [30, 45].

ЗАКЛЮЧЕНИЕ

Результаты проведенных исследований показали, что вышеописанная нейрональная поведенческая модель ориентации в радиальном лабиринте с дифференцированным подкреплением может быть успешно использована для изучения нейрональной организации, архитектуры и алгоритма поискового поведения испытуемых в пространстве, диагностики нарушений внимания и памяти [1, 3], конструирования роботизированных устройств, предназначенных для автономной ориентации в пространстве и целеполагания [28, 29, 35], а также в качестве основы для создания других автоматизированных устройств. Одним из возможных направлений использования описанной нейрональной модели адаптивного поискового поведения в радиальном лабиринте может быть моделирование и создание системы автоматического управления движением автомобилей в различных дорожных ситуациях: 1) при регулируемом светофором дорожном движении на перекрестке (на изолированное включение сенсорных сигналов светофора); 2) на нерегулируемом перекрестке (включены все сигналы светофора); 3) при поломке или отсутствии светофора (необходимость принятия решения автоматизированной системой о времени и порядке движения с учетом сенсорного анализа окружающей обстановки и правил приоритетного движения на нерегулируемом перекрестке).

Разработанная модель адаптивного поискового поведения в радиальном лабиринте с дифференцированным вознаграждением успешно применяется при тестировании как животных, так и человека с сохранением всех условий и процедуры проведения эксперимента [1–3]. Сходные тактики решения задач животными и человеком при изменении условий реагирования в радиальном лабиринте [3, 4, 17] указывают на перспективные возможности использования описанной модели подкрепляемого обучения при конструировании не только биоподобных, но и антропоморфных роботов.

Работа выполнена при финансовой поддержке Программы фундаментальных научных исследований государственных академий на 2014—2020 гг. (ГП-14, раздел 63).

СПИСОК ЛИТЕРАТУРЫ

[1] Альбертин С. В. Способ диагностики синдрома дефицита внимания в эксперименте на животных. Бюл. изобретений. 4(3): 331. 2006.

[2] Альбертин С. В. Автоматизированное устройство для тестирования внимания. Бюл. изобретений. 33: 920. 2009.

[3] Альбертин С. В. Диагностика синдрома дефицита внимания с гиперактивностью с использованием условнорефлекторного метода. Журн. неврологии и психиатрии им. С. С. Корсакова. 110(12): 56—60. 2010.

[4] Альбертин С. В. Модельные исследования центральной нервной системы. СПб. Дорн. 2011.

[5] Альбертин С. В. Участие гиппокампа и стриатума в формировании поискового поведения у ювенильных крыс: нейрофармакологический анализ. Междунар. научн. журн. «Инновационная наука». 8: 18—20. 2015.

[6] Альбертин С. В. Влияние стимуляции Доэргической системы мозга на пищевое предпочтение у крыс. Рос. физиол. журн. им. И. М. Сеченова. 102(10): 1137—1145. 2016.

[7] Альбертин С. В. Влияние режима условнорефлекторного переобучения крыс на поисковое поведение в радиальном лабиринте. Рос. физиол. журн. им. И. М. Сеченова. 102(11): 1302—1311. 2016.

[8] Альбертин С. В. Влияние сенсорных свойств подкрепления на формирование пищевого предпочтения у крыс. Сенсорные системы. 31(1): 3—8. 2017.

[9] Альбертин С. В. Влияние фрагментации зрительных навигационных сигналов на ориентацию крыс в радиальном лабиринте. Рос. физиол. журн. им. И. М. Сеченова. 103(8): 854—865. 2017.

[10] Альбертин С. В. Способ тестирования сенсомоторных реакций животных в условиях зрительного слежения. Сенсорные системы. 31(4): 290—296. 2017.

[11] Непомнящих В. А., Редько В. Г. Метод формирования поискового поведения. Искусств. интеллект и принятие решений. 1: 49—54. 2011.

[12] Редько В. Г. Модели адаптивного поведения — биологически инспирированный подход к искусственному интеллекту. Искусств. интеллект и принятие решений. 2: 11—23. 2008.

[13] Aguilar-Valles A., Sanchez E., de Gortari P., Balderas I., Ramirez-Amaya V., Bermudes-Rattoni F., Joseph-Bravo P. Analysis of stress response in rats trained in the water maze: differential expression of corticotropine-releasing hormone, glucocorticoid receptors and brain — derived neurotrophic factor in limbic regions. Neuroendocrinology. 82: 306—319. 2005.

[14] Albertin S. V., Mulder A. B., Tabuchi E., Zugaro M., Wiener S. I. Accumbens medial lesions impair performance in rats performing in a differently rewarded plus maze. Soc. Neurosci. Abstr. California. 24: 1909. 1998.

[15] Albertin S. V., Mulder A. B., Tabuchi E., Zugaro M., Wiener S. I. Lesions of the medial shell of the n. accumbens impair rats in finding larger rewards but spare reward-seeking behavior. Behav. Brain Res. 117: 173—183. 2000.

[16] Albertin S. V., Wiener S. I. Neuronal activity in the nucleus accumbens and hippocampus in rats during formation of seeking behavior in a radial maze. Bull. Exp. Biol. Med. 158: 405—409. 2015.

[17] Albertin S. V. Diagnosis of attention deficit hyperactivity disorder using a conditioned reflex approach. Neurosci. Behav. Physiol. 41(9): 906—910. 2011.

[18] Anderson M. C., Green C. Suppressing unwanted memories by executive control. Nature. 10(6826): 366—369. 2001.

[19] Arbib M., Metta G., van der Smagt P. Handbook of Robotics, chapter Neurorobotics: From vision to action. Berlin. Springer-Verlag. 1453—1480. 2008.

[20] Benchenane K., Peyrache A., Khamassi M., Tierney P. L., Gioanni Y., Battaglia F. P., Wiener S. I. Coherent theta oscillations and reorganization of spike timing in the hippocampal-prefrontal network upon learning. Neuron. 66: 921—936. 2010.

[21] Daw N. D., Niv Y., Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. 8: 1704—1711. 2005.

[22] Edelman G. M. Learning in and from Brain-Based Devices. Science. 318: 1103—1105. 2007.

[23] Fischer S., Diekelmann S., Born J. Sleep's role in the processing of unwanted memory. J. Sleep Res. 20(2): 267—270. 2011.

[24] Gentry R. N., Lee B., Roesch M. R. Phasic dopamine release in the rat n. accumbens predicts approach and avoidance performance. Nat. Commun. 7: 3154. 2016. doi: 10.1038/ncomms13154.

[25] Gluth S., Sommer T., Rieskamp J., Buchel C. Effective connectivity between hippocampus and ventromedial prefrontal cortex controls preferential choices from memory. Neuron. 86: 1078—1090. 2015.

- [26] *Harrison F. E., Hosseini A. H., McDonald M. P.* Endogenous stress responses in water maze and Barnes maze spatial memory tasks. *Behav. Brain Res.* 198(1): 247—251. 2009.
- [27] *Holscher C.* Stress impairs performance in spatial water maze learning tasks. *Behav. Brain Res.* 100: 225—235. 1999.
- [28] *Khamassi M., Lacheze L., Girard B., Berthoz A., Guillot A.* Actor-critic models of reinforcement learning in the basal ganglia: From natural to artificial rats. *Adaptive Behav.* 13(2): 131—148. 2005.
- [29] *Koene A., Prescott T. J.* Hippocampus, amygdala and basal ganglia based navigation control. *Lectures notes in computer science.* 5768LNCS. PART 1: 264—276. 2009.
- [30] *Kortenkamp D., Simmons R.* Robotic systems architectures and programming. In: B. Siciliano, O. Khatib (eds). *Handbook of Robotics.* Springer-Verlag. 187—206. 2008.
- [31] *Krichmar J. L., Nitz D. A., Gally J. A., Edelman G. M.* Characterizing functional hippocampal pathway in a brain-based device as it solves a spatial memory task. *PNAS.* 102(6): 2111—2116. 2005.
- [32] *Krichmar J. L., Seth A. K., Nitz D. A., Fleischer J. C., Edelman G. M.* Spatial navigation and causal analysis in a brain-based device modeling cortical-hippocampal interaction. *Neuroinformatics.* 3: 197—221. 2005.
- [33] *Lisman J. E., Grace A. A.* The hippocampal — VTA Loop: Controlling the entry of information into long-term memory. *Neuron.* 46: 703—713. 2005.
- [34] *Liu Z. W., Faraguna U., Cirelli C., Tononi G., Gao X.-B.* Direct evidence for wake-related increases and sleep related decreases in synaptic strength in rodent cortex. *J. Neurosci.* 23: 8671—8675. 2010.
- [35] *Meyer J.-A., Guillot A., Girard B., Khamassi M., Pirim P., Berthoz A.* The Psikharpax project: towards building an artificial rat. *Robotics Autonomous Syst.* 50: 211—223. 2005.
- [36] *Morris R.* Development of a water-maze procedure for studying spatial learning in the rats. *Neurosci. Meth.* 11: 47—60. 1984.
- [37] *Nicola S. M.* The flexible approach hypothesis: Unification of effort and cue responding hypotheses for the role of nucl. accumbens dopamine in the activation of reward-seeking behavior. *J. Neurosci.* 30(3): 16 585—16 600. 2010.
- [38] *Pennartz C. M., Lee E., Verheul J., Lipa P., Barnes C. A., McNaughton B. L.* The ventral striatum in off-line processing: ensemble reactivation during sleep and modulation by hippocampal ripples. *J. Neurosci.* 24: 6446—6456. 2004.
- [39] *Prescott T. J., Lepora N., Verschure P. F. M.* A future of living machines? International trends and prospects in biomimetic and biohybrid systems. *Proc. of SPIE — Internat. Soc. Optical Engineering.* 2014. DOI: 10.1117/12.2046.305.
- [40] *Ribeiro S.* Sleep and plasticity. *Pflügers Arch.* 463(1): 111—120. 2012.
- [41] *Rosenfeld C. S., Ferguson S. A.* Barnes Maze Testing strategies with small and large rodent models. *J. Vis. Exp.* 84: 51194. 2014. DOI: 10.3791/51194.
- [42] *Tabuchi E., Mulder A. B., Wiener S. I.* Position and behavioral modulation of synchronization of hippocampal and accumbens neurons discharges in freely moving rats. *Hippocampus.* 10: 718—728. 2000.
- [43] *Tseng K. Y., O'Donnell P.* DA/Glu interactions controlling prefrontal cortical pyramidal cell excitability involves multiple signaling mechanisms. *J. Neurosci.* 24: 5131—5139. 2004.
- [44] *Wiener S. I., Shibata R., Tabuchi E., Trullier O., Albertin S. V., Mulder A. B.* Spatial and behavioral correlates in nucleus accumbens neurons receiving hippocampal or prefrontal cortical inputs. *Int. Congr. Ser. Elsevier.* 1250: 275—292. 2003.
- [45] *Yarmohamadi M., Javadi H. H. C., Erfani H.* Improvement of robot path planning using particle swarm optimization in dynamic environments with mobile obstacle and targets. *Advanced Studies in Advanced Studies in Biology.* 3(1): 43—53. 2011.

Поступила 28 IV 2018
После доработки 8 VIII 2018