КОМПЬЮТЕРНЫЕ ИССЛЕДОВАНИЯ И МОДЕЛИРОВАНИЕ 2016 Т. 8 № 6 С. 941−950

Ки&М

АНАЛИЗ И МОДЕЛИРОВАНИЕ СЛОЖНЫХ ЖИВЫХ СИСТЕМ

УДК 004.94/004.81

Модель формирования первичных поведенческих паттернов с адаптивным поведением на основе использования комбинации случайного поиска и опыта

А. Г. Туманян[1,a], С. И. Барцев[1,2,b]

[1] Институт фундаментальной биологии и биотехнологии, Сибирского федерального университета, Россия, 660041, г. Красноярск, просп. Свободный, д. 79
[2] Институт биофизики СО РАН, Россия, 660036, г. Красноярск, Академгородок, 50/50

E-mail: [a] artur.tumanyan@gmail.com
E-mail: [b] bartsev@yandex.ru

Получено 21.10.2016, после доработки — 07.11.2016.
Принято к публикации 09.11.2016.

В работе предложен адаптивный алгоритм, моделирующий процесс формирования начальных поведенческих навыков на примере системы «глаза–манипулятор» анимата. Ситуация формирования начальных поведенческих навыков возникает, например, когда ребенок осваивает управление своими руками на основе понимания связи между исходно неидентифицированными пятнами на сетчатке своих глаз и положением реального предмета. Поскольку навыки управления телом не «вшиты» исходно в головной и спинной мозг на уровне инстинктов, то человеческому ребенку, как и большинству детенышей других млекопитающих, приходится осваивать эти навыки в режиме поискового поведения. Поисковое поведение начинается с метода проб и ошибок в чистом виде, затем его вклад постепенно уменьшается по мере освоения своего тела и окружающей среды. Поскольку образцов правильного поведения на этом этапе развития организм не имеет, то единственным способом выделения правильных навыков является положительное подкрепление при достижении цели. Ключевой особенностью предлагаемого алгоритма является фиксация в режиме импринтинга только завершающих действий, которые привели к успеху, или, что очень важно, привели к уже знакомой запечатленной ситуации, однозначно приводящей к успеху. Со временем непрерывная цепочка правильных действий удлиняется — максимально используется предыдущий позитивный опыт, а негативный «забывается» и не используется. Тем самым наблюдается постепенная замена случайного поиска целенаправленными действиями, что наблюдается и у реальных детенышей.

Тем самым алгоритм способен устанавливать соответствие между закономерностями окружающего мира и «внутренними ощущениями», внутренним состоянием самого анимата. В предлагаемой модели анимата использовалось 2 типа нейросетей: 1) нейросеть NET1, на вход которой подавались текущие положения кисти руки и целевой точки, а на выходе — двигательные команды, направляющие «кисть» манипулятора анимата к целевой точке; 2) нейросеть NET2, которая на входе получала координаты цели и текущей координаты «кисти», а на выходе формировала значение вероятности того, что анимату уже «знакома» эта ситуация и он «знает», как на нее реагировать. Благодаря такой архитектуре у анимата есть возможность опираться на «опыт» нейросети в распознанных ситуациях, когда отклик от сети NET2 близок к 1, и, с другой стороны, запускать случайный поиск, когда опыта функционирования в этой области зрительного поля у анимата нет (отклик NET2 близок к 0).

Ключевые слова: адаптивное поведение, поведенческая модель, анимат, нейронная сеть, самообучение

© 2016 Артур Геворкиевич Туманян, Сергей Игоревич Барцев

COMPUTER RESEARCH AND MODELING
2016 VOL. 8 NO. 6 P. 941−950

Ки&М

ANALYSIS AND MODELING OF COMPLEX LIVING SYSTEMS

UDC: 004.94/004.81

Model of formation of primary behavioral patterns with adaptive behavior based on the combination of random search and experience

A. G. Tumanyan[1], S. I. Bartsev[2]

[1] Siberian Federal University,
79 Svobodny st., Krasnoyarsk, 660041, Russia
[2] Institute of Biophysics,
50/50 Akademgorodok, Krasnoyarsk, 660036, Russia

Received 21.10.2016, after completion — 07.11.2016.
Accepted for publication 09.11.2016.

In this paper, we propose an adaptive algorithm that simulates the process of forming the initial behavioral skills on the example of the system 'eye-arm' animat. The situation is the formation of the initial behavioral skills occurs, for example, when a child masters the management of their hands by understanding the relationship between baseline unidentified spots on the retina of his eye and the position of the real object. Since the body control skills are not 'hardcoded' initially in the brain and the spinal cord at the level of instincts, the human child, like most young of other mammals, it is necessary to develop these skills in search behavior mode. Exploratory behavior begins with trial and error and then its contribution is gradually reduced as the development of the body and its environment. Since the correct behavior patterns at this stage of development of the organism does not exist for now, then the only way to select the right skills is a positive reinforcement to achieve the objective. A key feature of the proposed algorithm is to fix in the imprinting mode, only the final action that led to success, and that is very important, led to the familiar imprinted situation clearly leads to success. Over time, the continuous chain is lengthened right action — maximum use of previous positive experiences and negative 'forgotten' and not used. Thus there is the gradual replacement of the random search purposeful actions that observed in the real young.

Thus, the algorithm is able to establish a correspondence between the laws of the world and the 'inner feelings', the internal state of the animat. The proposed animat model was used 2 types of neural networks: 1) neural network NET1 to the input current which is fed to the position of the brush arms and the target point, and the output of motor commands, directing 'brush' manipulator animat to the target point; 2) neural network NET2 is received at the input of target coordinates and the current coordinates of the 'brush' and the output value is formed likelihood that the animat already 'know' this situation, and he 'knows' how to react to it. With this architecture at the animat has to rely on the 'experience' of neural networks to recognize situations where the response from NET2 network of close to 1, and on the other hand, run a random search, when the experience of functioning in this area of the visual field in animat not (response NET2 close to 0).

Keywords: adaptive behavior, behavioral model, animat, neural network, self-study

Citation: *Computer Research and Modeling*, 2016, vol. 8, no. 6, pp. 941–950 (Russian).

© 2016 Artur G. Tumanyan, Sergey I. Bartsev

1. Введение

Живые организмы, существуя в природе, руководствуются разного рода поведенческими навыками или паттернами для выживания: если организм голодает, значит нужно искать пищу, если рядом хищник, значит нужно убегать и др. Большинство паттернов — приобретенные в процессе жизни организма. Способность к формированию новых поведенческих паттернов помогает приспособиться к меняющейся среде и позволяет выживать в новых условиях. То, как устроены механизмы такого формирования, как они взаимодействуют друг с другом, важные и актуальные вопросы когнитивной биологии и искусственного интеллекта.

В настоящее время существует ряд биологически реалистичных методов обучения аниматов [Барцев, 1989; Редько, 2007; Редько, 2008; Carpenter, 1990; Горбань, 1990; Николлс, 2003; Гаазе-Рапопорт, 1987], т. е. роботов или их компьютерных моделей, предназначенных для изучения поведения живых организмов. Среди таких методов наибольшее распространение получили генетические алгоритмы [Гладков, 2006], спайковые нейронные сети [Синявский, 2011] и обучение с подкреплением (reinforcement learning — RL) [Sutton, 1998]. Эти методы хорошо показывают себя в ситуациях с небольшими размерностями входных данных, особенно метод обучения с подкреплением. Перечисленные выше алгоритмы функционируют без шаблона, задающего правильное (эффективное) поведение, при этом исследуется всё пространство состояний и время поиска удачного состояния не детерминировано. При больших размерностях алгоритмы RL теряют практическую ценность из-за сложности вычислений, которая растет как степенная функция. Одним из решений является использование метода редукции размерности обрабатываемых данных, hierarchical SFA network [Legenstein, 1998], но и данный метод не снимает все ограничения RL — приходится также изучать все пространство состояний, чтобы получить приемлемый результат.

Супервизорные алгоритмы [Rumelhart, 1986; Галушкин, 1974], широко используемые в нейрокомпьютеринге, могут обучаться лишь на основе задачника, содержащего информацию о правильной реакции анимата на конкретное воздействие, и поэтому необходимо наличие внешнего «учителя», чтобы составить всеобъемлющий задачник с правильными поведенческими паттернами. Другая отрицательная черта супервизорных алгоритмов — необходимость многократного предъявления данных, чтобы все весовые коэффициенты успели подстроиться (если это вообще возможно для выбранной задачи и выбранного метода подстройки коэффициентов).

В работе [Туманян, Барцев, 2014] была разработана модель импринта [Балацкий, 2007], благодаря которой анимат был способен в результате однократного предъявления образца ситуации запечатлеть последовательность действий, которые приводили к успеху. Дополнительно был обнаружен эффект сглаживания траектории, что указывает на формирование обобщений даже при однократном предъявлении ситуации. Особенностью модели является способность к запечатлению удачной последовательности действий в ответ на единственное критическое событие, что характерно для достаточно высокоорганизованных организмов. Так, животное, которое смогло выжить в экстремальной ситуации, в будущем, при повторении этой ситуации, будет выполнять именно ту последовательность действий, которая помогла ему избежать смерти.

При этом формирование траектории происходило на основе критерия близости изображений цели и «кисти» манипулятора на сетчатке глаза анимата, т. е. направление движения виртуального манипулятора выбиралось случайно, но если угловое расстояние между проекциями кисти и изображения цели возрастало, то происходило случайное переключение направления движения манипулятора. Это было упрощение, позволяющее сконцентрироваться на главной задаче — проверке возможности формирования хватательного навыка при единственном предъявлении, характерном для импринта. После того как это было показано, возник вопрос о том, как организмом может формироваться критерий достижения цели манипулятором, введенный в описанной работе «усилием воли» извне.

В данной работе главная цель — создание биологически реалистичного алгоритма обучения анимата, который способен устанавливать соответствие между закономерностями окру-

жающего мира (взаимное положение цели и «кисти» манипулятора) и «внутренними ощущениями» (положением исходно не идентифицированных пятен на сетчатке), приводящими к адекватной активной реакции анимата — захвату предмета. Особо отметим, что это не просто формирование условного рефлекса, а именно усвоение фундаментальных закономерностей отображения внешнего мира во внутреннее представление. Крайне желательно, чтобы этот алгоритм обучения анимата допускал совмещение процессов обучения с процессами функционирования и жизнедеятельности.

2. Описание эксперимента

Для разработки алгоритма использовались виртуальные стенды, разработанные в среде Lazarus [Туманян, Барцев 2014] и на базе языка Python, в которых моделируется функционирование анимата. Стенд, на котором проводился эксперимент по моделированию анимата (рис. 1), имел 2 одномерные сетчатки, на которые проецировалась кисть манипулятора и целевая точка, а также сама рука-манипулятор с двумя суставами (2 степени свободы). Для простоты реализации в текущем стенде использовались декартовы координаты целевых точек и кисти руки $x = \text{xpos}/W$, $y = \text{ypos}/H$, где xpos и ypos — положение на графическом поле стенда, W и H — ширина и высота поля. Фактически мы предполагаем наличие препроцессора, который переводит бинокулярную информацию об углах проекций объектов на сетчатках «глаз» в декартовы координаты этих объектов. Виртуальный стенд был способен отображать траектории движения руки-манипулятора, накапливать фрагменты траекторий и генерировать целевую точку в случайном месте на зрительном поле в процессе обучения.

Процесс обучения нейросетей осуществляется по успешным фрагментам траекторий, выступающим в качестве обучающего паттерна для алгоритма обратного распространения ошибки, реализованного в библиотеке PyBrain. Важно отметить, что выбор именно этого алгоритма не существенен, в принципе возможно использование любого супервизорного алгоритма. Визуализация была сделана с помощью пакета PyGame.

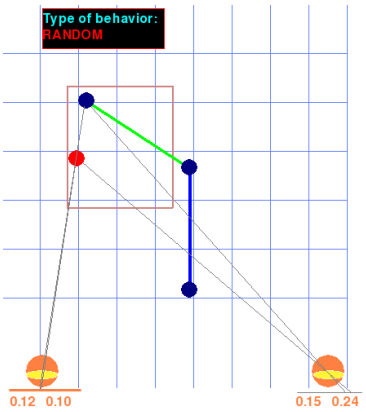


Рис. 1. Желтым показаны «глаза» анимата, на которые происходит проецирование; целевая точка отмечена красным цветом, красным прямоугольником обозначено зрительное поле, в котором появляются целевые точки; оранжевыми цифрами показаны текущие декартовы координаты для целевой точки и руки

Функционирование анимата — движение руки — происходит по тактам, в каждый такт каждый из шарниров может повернуться на (–1, 0, 1) градуса. Углы поворота при случайном поиске задаются генератором случайных чисел, а при функционировании — обученной нейронной сетью. Обучение организовано следующим образом (рис. 2): случайным образом формируется целевая точка, и анимат начинает делать случайные шаги. В случае когда анимат достигает цели, фрагмент траектории, непосредственно предшествовавший этому событию, помещается в задачник для нейросети NET1 (таблица 1), которая управляет движениями. Кроме то-

го, данные о движении руки также помещаются в другой задачник для обучения второй нейросети NET2, отвечающей за распознавание ситуации (таблица 2).

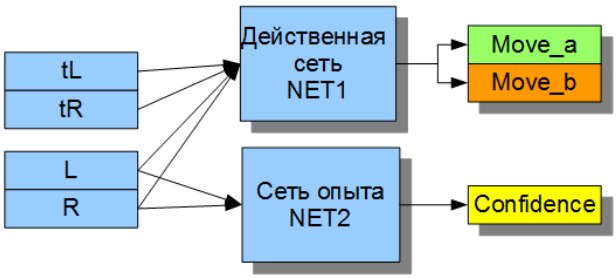


Рис. 2. Архитектура когнитивной системы анимата. Нейросеть NET1 отвечает за формирования двигательного паттерна, который способен приводить анимат к целевой точке. Нейросеть NET2 отвечает за запоминания ситуаций, в которых анимат уже находился, а значит, имеет опыт достижения цели

После того как данные траектории помещены в задачники, нейросети обучаются. Когда начинается новая итерация для той же целевой точки, но другого начального положения руки, анимат начинает случайные блуждания до тех пор, пока не встретит целевую точку либо же не столкнется с «сохраненными» в сети NET2 фрагментами успешных траекторией. То есть в случае, когда нейросеть NET2 выдает на выходе значение уверенности распознавания ситуации близкое к 1, случайный поиск прекращается и управление передается сети NET1, которая хранит в своей структуре весовых коэффициентов правильную модель действий, для «знакомых» ситуаций. Важно, что успешной считается не только траектория, которая завершилась захватом цели, но также и траектория, которая «вышла» в окрестности «знакомых» фрагментов траекторий, усвоенных обеими нейросетями. Тем самым финальный фрагмент такой траектории тоже добавляется в задачник и используется для дальнейшего обучения нейросетей.

Обучение здесь неотделимо от процесса функционирования, так как подстройка весовых коэффициентов происходит при каждом успешном раунде нахождения цели. Каждая новая успешная траектория приводит к тому, что фрагмент траектории добавляется к общему задачнику и сеть доучивается. Фрагмент задачника для нейросети NET1 представлен в таблице 1. Задачник имеет традиционную для супервизорных алгоритмов структуру: строка задачника содержит величины входных сигналов, поступающих от «рецепторов» (L, R, tL, tR) и требуемый от нейросети отклик на выходах (α, β).

Таблица 1. Задачник для нейросети NET1 анимата

№	L	R	tL	tR	α	β
1	0.49	0.04	0.67	0.30	0	–1
2	0.48	0.04	0.67	0.30	–1	1
3	0.52	0,07	0.67	0.30	–1	1
4	0.54	0.10	0.67	0.30	–1	0
5	0.57	0.14	0.67	0.30	–1	0
6	0.59	0.17	0.67	0.30	–1	0
7	0.61	0.21	0.67	0.30	–1	1
8	0.62	0.25	0.67	0.30	–1	1

Примечание. В столбцах: L, R — координаты кисти руки-манипулятора; tL, tR — координаты целевой точки; α и β — двигательные команды.

Фрагмент задачника для нейросети NET2 представлен в таблице 2, M — логическое значение, которое указывает на принадлежность точки, в которой на данный момент находится рука анимата, к какой-либо сохраненной успешной траектории, т. е. на вход нейросети подают-

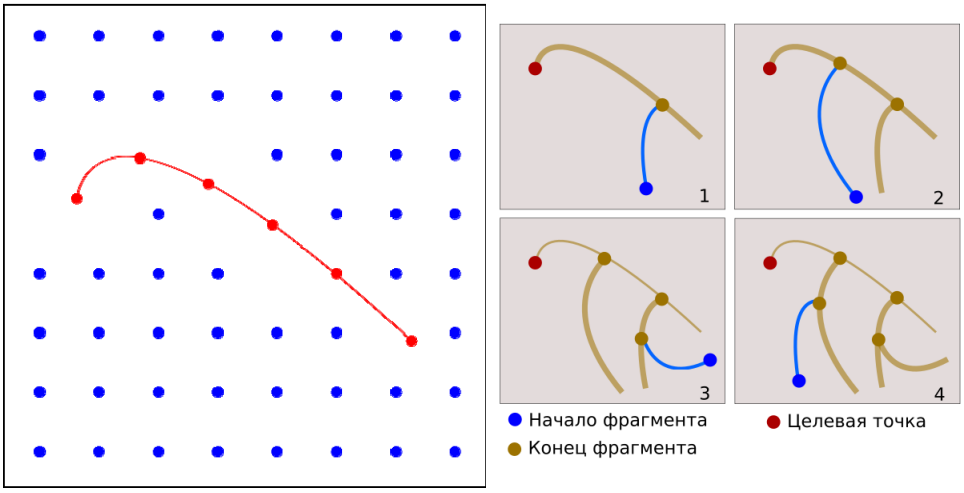
ся все точки из зрительного поля, принадлежащие успешным траекториям ($M = 1$), и точки, им не принадлежащие ($M = 0$).

Таблица 2. Задачник для нейросети NET2 анимата

№	L	R	M
1	0.49	0.04	1
2	0.48	0.04	1
3	0.52	0,07	0

Примечание. Столбцы L, R — координаты кисти руки-манипулятора, M — маркер, которые указывает на принадлежность точки к успешной траектории.

Для уменьшения размерности задачника (рис. 3, *а*) «пустые» точки, т. е. те, через которые не проходили траектории ($M = 0$), были «прорежены» для обеспечения примерного паритета количества точек зрительного поля, соответствующих фрагментам правильных траекторий и незнакомым еще областям. При этом размерность входного вектора была уменьшена с ~10000 до ~100 элементов. Показанный способ облегчил процесс обучения нейросети и помог ей запоминать любую комбинацию траекторий и пустых областей.



(а) (б)

Рис. 3. (а) Уменьшение размерности обучающей выборки; (б) схематическое изображение процесса накопления опыта аниматом

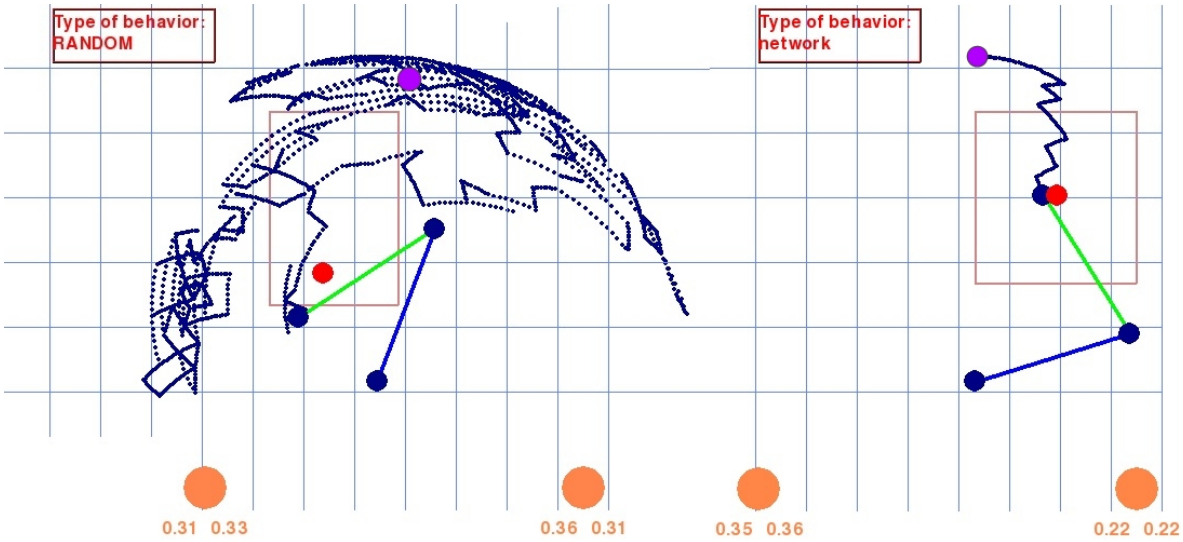
Процесс обучения анимата (рис. 3, *б*) проходил итеративно, и каждая новая захваченная траектория присоединялась к уже существующей сети траекторий, которая вела к целевой точке. Анимату было достаточно попасть в зону узнавания, а далее уже перехватывала управление действенная сеть, которая приводила к целевой точке.

На рис. 3, *б* изображение 1 иллюстрируется как анимат; совершая новое случайное движение (синяя траектория), он попадает в область, в которой уже бывал (коричневая траектория); нейросеть, отвечающая за опыт, выдает отклик, близкий к 1, и далее включается действенная сеть, которая приводит анимата к успеху. При этом финальный фрагмент синей траектории на изображении 1 становится частью опыта анимата и далее пользуется повсеместно, при каждом запуске (изображения 2–4).

Тем самым для увеличения скорости обучения фиксировались только последние действия, которые привели к успеху; со временем непрерывная цепочка правильных действий удлиняется — максимально используется предыдущий позитивный опыт, а негативный «забывается» и не используется. Этот способ может применяться для освоения сложных временны́х паттернов.

3. Результаты

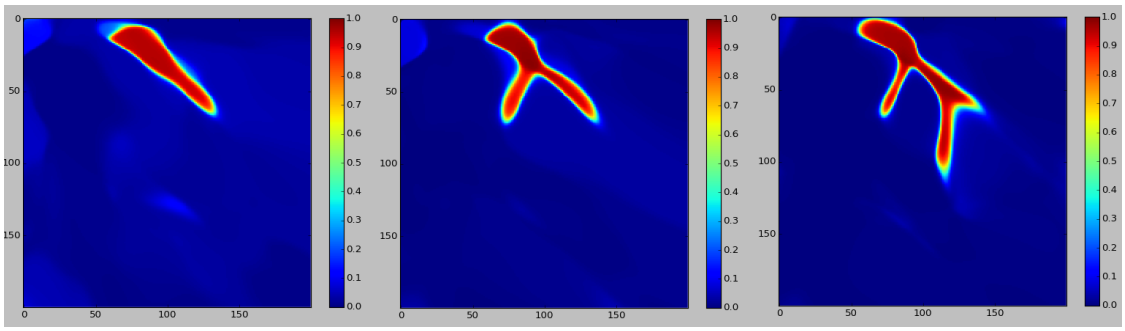
В начале своей работы анимат совершал исключительно случайные действия (рис. 4, *а*), ожидаемо затрачивая много времени на поиск цели. Затем, по мере накопления успешных завершающих фрагментов траекторий в его задачнике, обучения сети на их основе, анимат начинал пользоваться «опытом» нейросети (рис. 4, *б*).



(а) (б)

Рис. 4. (а) Демонстрация работы стенда в режиме чистого поискового поведения, когда рука-манипулятор случайно «блуждает» в поисках цели и (б) рука функционирует, основываясь на опыте нейросети. Начала траекторий отмечены фиолетовыми кругами, пунктирные отображения траекторий не отличаются от сплошных по назначению, а служат лишь для экономии вычислительных ресурсов при формировании задачников и обучения

Оценку эффективности обучения нейросети NET2 (определяющей место) проводили визуально. После каждой успешной итерации обучения полученная нейросеть подвергалась «сканированию», т. е. последовательно для каждой точки зрительного поля вычислялся отклик нейросети (рис. 5).



(а) (б) (в)

Рис. 5. Карта памяти сети анимата: (а) карта для единичной траектории; (б) карта для двух сохраненных траекторий; (в) карта для трех траекторий

Сканирование позволило заглянуть внутрь процесса «запоминания» места опытной нейросетью анимата. Рис. 5, *а* иллюстрирует карту памяти сети NET2, которая содержит знание только об одной траектории, которая приводит к целевой точке; на рис. 5, *б* в карте памяти сети

появляется новая траектория, почти перпендикулярная старой. Попадание анимата в любую точку области с высоким откликом (>0.7) передает управления действенной сети. С каждой новой успешной траекторией увеличивается область гарантированного распознавания ситуации. Фрагмент траектории, который предшествовал попаданию в зону узнавания, будет сохранен и в дальнейшем использован. Со временем цепочка правильных действий удлиняется, увеличивается «опыт» анимата, что соответствует расширению области «узнавания» на зрительном поле.

Мерой эффективности обучения являлось время поиска цели аниматом. Для каждой отдельно сформированной цели время поиска различное; поскольку это случайный поиск, то оно может различаться в разы. Если алгоритм обучения эффективен, то в ходе обучения время поиска целевой точки должно уменьшаться вследствие увеличения влияния на поведение анимата отклика «опытной» нейросети. Мы последовательно генерировали точки на зрительном поле и замеряли время поиска (время между попаданием руки в целевые точки). На рис. 6 показаны диаграммы временны́х затрат для анимата: в случае случайного поиска (6, *а*) и в случае комбинированного поиска (6, *б*), сочетающего случайный поиск при незнакомом паттерне (ситуации) и переключение на функционирование нейросети при распознавании знакомого паттерна (ситуации). Существенные различия в среднем времени поиска цели очевидны и не требуют специальной статистической обработки.

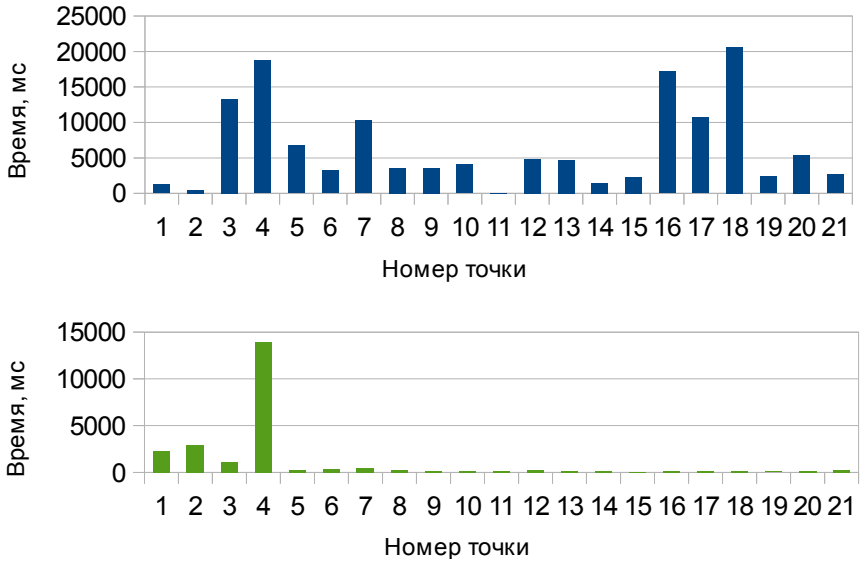


Рис. 6. Время поиска цели в секундах в зависимости от номера цели (21 целевая точка): синяя диаграмма (а) — только случайный поиск, зеленая диаграмма (б) — комбинация случайного поиска и отклика обученной нейросети

Алгоритм удовлетворяет важному с точки зрения моделирования поведения животных условию: он способен улучшаться в процессе функционирования, и улучшения эти носят последовательный характер.

4. Заключение

В статье разработан биологически реалистичный алгоритм обучения анимата, ключевой особенностью которого является то, что фиксируются только последние действия, которые привели к успеху, и со временем непрерывная цепочка правильных действий удлиняется: максимально используется предыдущий позитивный опыт, а негативный «забывается» и не используется.

Другая особенность алгоритма заключается в том, что процессы обучения происходят неразрывно от процессов функционирования и «жизнедеятельности», т. е. анимат с самого начала

может комбинировать случайные действия и действия, основанные на опыте. Комбинируя обучение на основе фрагментов траекторий и случайный поиск, анимат имеет возможность, во-первых, оптимизировать свое поведение, сделать траектории короче, движения — плавными [Туманян, 2014], а во-вторых, анимат всегда улучшает свои характеристики в процессе жизнедеятельности благодаря доучиванию сети при каждом успешном исходе.

Разработанный метод хорошо соотносится с поведением живых организмов: когда организм не знает, как поступить, он пробует случайные движения; после нескольких попаданий в цель его способность резко возрастает, время достижения цели соответственно падает.

Список литературы

Балацкий Е. В. Вспомагательные импринты и поведение человека // Вестник Российской академии наук. — 2007. — Т. 77, № 10.

Balatsky E. V. Accessory imprints and behavior // Herald of the Russian Academy of Sciences. — 2007. — Vol. 77, No. 10. (Original Russian paper: *Balackij E. V.* Vspomagatel'nye imprinty i povedenie cheloveka // Vestnik Rossijskoj Akademii Nauk. — 2007. — T. 77, No. 10.)

Барцев С. И., Гилев С. Е., Охонин В. А. Принцип двойственности в организации адаптивных сетей обработки информации // Динамика химических и биологических систем. — Новосибирск: Наука, 1989. — С. 6–55.

Bartsev S. I., Gilev S. E., Okhonin V. A. duality principle in the organization of the adaptive information processing networks // Dynamics of chemical and biological systems. — Novosibirsk: Science, 1989. — P. 6–55. (Original Russian paper: *Barcev S. I., Gilev S. E., Ohonin V. A.* Princip dvojstvennosti v organizacii adaptivnyh setej obrabotki informacii // Dinamika himicheskih i biologicheskih sistem. — Novosibirsk: Nauka, 1989. — S. 6–55.)

Гаазе-Рапопорт М. Г., Поспелов Д. А. От амебы до робота: модели поведения. — М.: Наука; Физматлит, 1987. — 288 с.

Haase-Rapoport M. G., Pospelov D. A. From amoeba to the robot: behaviors. — M.: Nauka; Fizmatlit, 1987. — 288 c. (Original Russian paper: *Gaaze-Rapoport M. G., Pospelov D. A.* Ot ameby do robota: modeli povedenija. — M.: Nauka; Fizmatlit, 1987. — 288 s.)

Галушкин А. И. Синтез многослойных систем распознавания образов. — М.: Энергия, 1974.

Galushkin A. I. Synthesis of multilayer pattern recognition systems. — M.: Energy, 1974. (Original Russian paper: *Galushkin A. I.* Sintez mnogoslojnyh sistem raspoznavanija obrazov. — M.: Jenergija, 1974.)

Гладков Л. А., Курейчик В. В., Курейчик В. М. Генетические алгоритмы: Учебное пособие. — 2-е изд. — М.: Физматлит, 2006. — С. 320.

Gladkov L. A., Kureichik V. V., Kureichik V. M. Genetic Algorithms: Tutorial. — 2nd ed. — M.: Fizmatlit, 2006. — P. 320. (Original Russian paper: *Gladkov L. A., Kurejchik V. V., Kurejchik V. M.* Geneticheskie algoritmy: Uchebnoe posobie. — 2-e izd. — M.: Fizmatlit, 2006. — S. 320.)

Горбань А. Н. Обучение нейронных сетей. — М.: СССР–США СП «ПараГраф», 1990. — 160 с.

Gorban A. N. Training neural networks. — M.: USSR–US joint venture “ParaGraph”, 1990. — 160 p. (Original Russian paper: *Gorban' A. N.* Obuchenie nejronnyh setej. — M.: SSSR–SShA SP “ParaGraf”, 1990. — 160 s.)

Николлс Д., Мартин Р., Валлас Б., Фукс П. От нейрона к мозгу / Пер. П. М. Балабана, А. В. Галкина, Р. А. Гиниатуллина, Р. Н. Хазипова, Л. С. Хируга. — М.: Едиториал УРСС, 2003. — 672 с.

Nicholls D., Martin R., Wallace B., Fuchs P. From the neuron to the brain / Trans. P. M, Balaban, A. V. Galkin, R. A. Giniatullin, R. N. Hazipova, L. S. Reconstructive surgery. — M.: Editorial URSS, 2003. — 672 p. (Original Russian paper: *Nikolls D., Martin R., Vallas B., Fuks P.* Ot nejrona k mozgu / Per. P. M. Balabana, A. V. Galkina, R. A. Giniatullina, R. N. Hazipova, L. S. Hiruga. — M.: Editorial URSS, 2003. — 672 s.)

Редько В. Г. Модели адаптивного поведения — биологически инспирированный подход к искусственному интеллекту // Искусственный интеллект и принятие решений. —2008. — Т. 2.

Redko V. G. Models of adaptive behavior — biologically inspired approach to artificial intelligence // Artificial intelligence and decision-making. — 2008. — Vol. 2. (Original Russian paper: Red'ko V. G. Modeli adaptivnogo povedenija — biologicheski inspirirovannyj podhod k iskusstvennomu intellektu // Iskusstvennyj intellekt i prinjatie reshenij. — 2008. — T. 2.)

Редько В. Г. Перспективы исследований на стыке информатики и биологии // Нейроинформатика. — 2007. — Т. 2, № 1.

Redko V. G. Prospects for research at the intersection of computer science and biology // Neuroinformatics. — 2007. — Vol. 2, No. 1. (Original Russian paper: *Red'ko V.G.* Perspektivy issledovanij na styke informatiki i biologii // Nejroinformatika. — 2007. — T. 2, No. 1.)

Синявский О. Ю., Кобрин А. И., Обучение с подкреплением спайковой нейронной сети в задаче управления агентом в дискретной виртуальной среде // Нелинейная динамика. — 2011. — Т. 7, № 4.

Sinyavsky O. J., Kobrin A. I. Reinforcement learning spiking neural network in the task management agent in discrete virtual environment // Nonlinear Dynamics. — 2011. — T. 7, № 4. (Original Russian paper: *Sinyavsky O. J., Kobrin A. I.* Obuchenie s podkrepleniem spajkovoj nejronnoj seti v zadache upravlenija agentom v diskretnoj virtual'noj srede // Nelinejnaja dinamika. — 2011. — Vol. 7, No. 4.)

Туманян А. Г., Барцев С. И. Простейшая поведенческая модель формирования импринта // Компьютерные исследования и моделирование. — 2014. — Т. 6, № 5. — С. 793–802.

Tumanyan A. G., Bartsev S. I. Prostejshaja povedencheskaja model' formirovanija imprinta [The simplest model of behavioral imprint] // Computer studies and modeling. — 2014. — Vol. 6, No. 5. — P. 793–802 (in Russian).

Туманян А. Г., Барцев С. И. Свидетельство о государственной регистрации программы для ЭВМ № 2016611753, «Программа для создания реалистичной модели простейшего двигательного паттерна мобильного робота».

Tumanyan A. G., Bartsev S. I. The simplest model of behavioral imprint. Computer simulation studies. — 2014. — Vol. 6, No. 5. — P. 793–802. (Original Russian paper: *Tumanjan A. G., Barcev S. I.* Svidetel'stvo o gosudarstvennoj registracii programmy dlja JeVM № 2016611753, «Programma dlja sozdanija realistichnoj modeli prostejshego dvigatel'nogo patterna mobil'nogo robota».)

Carpenter G. A., Grossberg S. ART 3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures // Neural Networks. — 1990. — Vol. 3. — P. 129–152.

Legenstein R., Wilbert N., Wiskott L. Reinforcement Learning on Slow Features of High-Dimensional Input Streams // PLoS Comput Biol. — 2010. — Vol. 8, e1000894. doi:10.1371/journal.pcbi.1000894.

Rumelhart D. E., Hinton G. E.,Williams R. J. Learning internal representation by error propagation // Paraller Distributed Processing. — 1986. — Vol. 1, No. 8.

Sutton R. S., Barto A. G. Reinforcement learning: An introduction. — Cambridge, MA: MIT Press. — 1998. — 323 p.