

УДК: 57.02

Простейшая поведенческая модель формирования импринта

А. Г. Туманян^{1,a}, С. И. Барцев^{1,2,b}

¹ Институт фундаментальной биологии и биотехнологии
Сибирского федерального университета,
Россия, 660041, г. Красноярск, Свободный просп., д. 79

² Институт биофизики СО РАН,
Россия, 660036, г. Красноярск, Академгородок, 50/50

E-mail: ^a artur.tumanyan@gmail.com, ^b bartsev@yandex.ru

Получено 18 сентября 2014 г.,
после доработки 20 октября 2014 г.

Формирование адекватных поведенческих паттернов в условиях неизвестного окружения осуществляется через поисковое поведение. При этом быстрее формирование приемлемого паттерна представляется более предпочтительным, чем долгая выработка совершенного паттерна, через многократное воспроизведение обучающей ситуации. В экстремальных ситуациях наблюдается явление импринтирования — мгновенного запечатления поведенческого паттерна, обеспечившего выживание особи. В данной работе предложены гипотеза и модель импринта, когда обученная по единственному успешному поведенческому паттерну нейронная сеть анимата демонстрирует эффективное функционирование. Реалистичность модели оценена путем проверки устойчивости воспроизведения поведенческого паттерна к возмущениям ситуации запуска импринта.

Ключевые слова: импринт, поведенческая модель, анимат, нейронная сеть, самообучение

Simple behavioral model of imprint formation

A. G. Tumanyan¹, S. I. Bartsev^{1,2}

¹ Siberian Federal University, 79 Svobodny St., Krasnoyarsk, 660041, Russia

² Institute of Biophysics, 50/50 Akademgorodok, Krasnoyarsk, 660036, Russia

Abstract. — Formation of adequate behavioral patterns in condition of the unknown environment carried out through exploratory behavior. At the same time the rapid formation of an acceptable pattern is more preferable than a long elaboration perfect pattern through repeat play learning situation. In extreme situations, phenomenon of imprinting is observed — instant imprinting of behavior pattern, which ensure the survival of individuals. In this paper we propose a hypothesis and imprint model when trained on a single successful pattern of virtual robot's neural network demonstrates the effective functioning. Realism of the model is estimated by checking the stability of playback behavior pattern to perturbations situation imprint run.

Keywords: imprint, behavioral model, animat, a neural network, self-study

Citation: *Computer Research and Modeling*, 2014, vol. 6, no. 5, pp. 793–802 (Russian).

Введение

Общеизвестно, что детеныши высокоорганизованных животных, как правило, не имеют от рождения развернутых поведенческих паттернов, обеспечивающих им адекватное реагирование на внешние воздействия. Можно предположить, что это свойство влечет за собой как отрицательные последствия (требуется время на освоение окружения, то есть на формирование адекватных поведенческих паттернов), так и положительные (животное способно приспособиться к новой среде, с которой вид прежде не сталкивался). Стремление понять механизмы обучения или адаптации к изменчивой, динамической среде вызвало появление бионического направления, представленного как коннекционистскими моделями — нейронными сетями [Гаврилов, 2000; Коршунов, 1972; Carpenter, 1990], так и алгоритмами эволюционного моделирования [Тарасов, 1999].

Очевидно, что период формирования начальных (первичных) поведенческих паттернов [Гаазе-Рапопорт, 1987] является весьма критичным по отношению к последующей жизни особи, поскольку дальнейшее освоение окружающего мира будет на них опираться и от того, насколько эти первичные паттерны эффективны, будет зависеть эффективность поведения зрелой особи. Налицо некоторое противоречие: с одной стороны, формирование первичных (и вероятнее всего, базовых) поведенческих паттернов должно происходить быстро, а с другой стороны, они должны быть достаточно эффективны. Отсюда следует, что режим тщательного «вылизывания» поведенческих навыков (рефлексов), как это делают, например, спортсмены, детенышу не подходит — необходимо как можно быстрее освоить («схватить») навык на уровне приемлемой результативности (субоптимальный навык). Такое быстрое (за одно предъявление) формирование поведенческого навыка в природе известно и называется импринтированием, или импринтом [Балацкий, 2007; Баранова, 2009]. Этим термином называют как запечатление первого увиденного подвижного объекта в качестве образа матери, так и запечатление (фиксация) определенной последовательности действий в экстремальной обстановке, то есть на фоне сильного эмоционального переживания. Впервые в 1935 году Конрад Лоренц стал использовать термин «импринт» для описания поведения серых гусей в известных сериях экспериментов [Лоренц, 1997].

В основу предлагаемой модели формирования импринта положена гипотеза, что он формируется на основе единичного успешного поведенческого акта, который является результатом случайного или почти случайного поиска, при этом сопровождающее импринт сильное эмоциональное переживание интенсифицирует формирование межнейронных связей в мозгу, запечатлевающих этот акт.

Однако запечатлевается конкретная ситуация, в которой присутствуют многочисленные детали, что потенциально может привести к невозможности повторного запуска запечатленного поведенческого акта просто вследствие малой вероятности точного воспроизведения ситуации, при которой произошло формирование импринта. Поэтому для того, чтобы предлагаемая гипотеза могла претендовать на реалистичность, необходимо ее дополнить предположением, что в процессе формирования импринта происходит обобщение закономерностей его внутренней структуры, приводящее к грубости (робастности) запечатленного паттерна. Эта робастность приводит к тому, что для запуска поведенческого акта не требуется, чтобы текущая ситуация в точности соответствовала исходной — достаточно их подобия.

Понимание на уровне воспроизведения механизмов импринтирования будет полезно для этологов, а также в контексте создания автономных роботов с гибким поведением, так называемых аниматов.

В настоящее время существует ряд биологически реалистичных методов обучения аниматов, которые часто применяются в моделировании: генетические алгоритмы [Гладков, 2006], обучение спайковых нейронных сетей [Синявский, 2011] и обучение с подкреплением [Хайкин, 2006; Рассел, 2006; Sutton, 1998]. Метод обучения с подкреплением заслуживает особого внимания поскольку он применим в условиях заранее незнакомого окружения. Идея данного под-

хода заключается в том, что за каждое совершенное действие анимат получает подкрепление — некоторое действительное число, которое может быть положительным (награда) или отрицательным (наказание). Задачей анимата является максимизация суммарной награды, которую он может получить с течением времени, подстраивая свое поведение таким образом. Формирование поведенческого навыка происходит благодаря модификации нейронных сетей [Барцев, 1989; Rumelhart, 1986] — так называемого «критика» и «контроллера» [Редько, 2007; Жданов, 2010; Демин, 2008], которые являются системами внутренней оценки качества действий анимата и его функционирования.

Однако вышеперечисленные методы соответствуют режиму постепенного, достаточно длительного формирования навыка, требуют большого количества входных данных и не подходят для описания импринта. На данный момент не существует моделей импринтинга, которые бы описывали процесс «запечатления» двигательного паттерна.

Целью работы является разработка модели формирования импринта и оценка ее физиологичности (реалистичности).

Материалы и методы

Для проверки вышеприведенной гипотезы разработана вычислительная среда (рис. 1) — виртуальный стенд, который включает в себя графический интерфейс, модуль зрительного препроцессора, модуль квази-случайного поиска и нейросетевой модуль. В качестве моделируемого поведенческого паттерна рассматривается формирование хватательного навыка в предельно упрощенной виртуальной 2D-среде. Вычислительная среда служит лишь для проверки функционирования модели импринтинга и является второстепенной разработкой по отношению к сути работы — модели обучения при однократном предъявлении (модели импринтинга).

Модель руки-манипулятора, упрощенный вариант многозвенного манипулятора [Кривошеев, 2012], состоит из двух сочленений, a и b , которые могут поворачиваться вокруг шарниров, один из которых («плечевой») неподвижен и закреплен на уровне «глаз» (рис. 1), а второй шарнир — «локтевой». На плоскости располагаются две разноцветные точки: одна представляет собой объект, а вторая — конечную точку «руки», кисть. Информация о текущем положении целевой точки и конечной точки руки проецируется на одномерные «сетчатки» левого и правого глаза.

Искусственная нейронная сеть — это математическая модель, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма. Нейросеть представляет собой систему соединенных и взаимодействующих между собой искусственных нейронов. С математической точки зрения обучение нейронных сетей — это многопараметрическая задача нелинейной оптимизации. Нейронные сети не программируются в привычном смысле этого слова, а обучаются. Возможность обучения — одно из главных преимуществ нейронных сетей перед традиционными алгоритмами, и технически обучение заключается в нахождении коэффициентов связей между нейронами. В процессе обучения нейросеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. Все задачи связанные с обучением и тестированием нейронной сети, производились с помощью нейросетевой библиотеки при помощи подключения .DLL файла к стенду анимата.

Для упрощения обработки зрительной информации нейронной сетью использованы знания о том, что в зрительном анализаторе животных существуют определенные группы нейронов, реагирующие на чередующиеся горизонтальные и вертикальные полосы [Смит, 2005]. Они служат для извлечения инвариантной информации из зрительного потока и уменьшения нагрузки на мозг. В данной работе предполагается наличие некоторого препроцессора, который работает аналогично клеткам стриарной коры [Николлс, 2003], т. е. преобразует угловые состояния объекта и руки, поступающие с «сетчатки», в их декартовы координаты.

Имеется четыре режима работы стенда: 1) выбор положения объекта вручную или с помощью генератора случайных чисел; 2) квазислучайный поиск такого положения «руки» при котором «кисть» совмещена с объектом, при этом происходит потактовая запись положения «кисти» и объекта (входов) и команд поворота шарниров (выходов); 3) обучение нейронной сети с помощью пакета FANN [URL: www.leenissen.dk/fann/wp], где в качестве обучающей выборки используется потактовая запись значений входов и выходов во время случайного поиска; 4) «захват» «рукой» объекта, осуществляемый под управлением обученной нейронной сети, — функционирование.

«Рука» движется по тактам, в каждый такт каждый из шарниров может повернуться на $(-3, -2, -1, 0, 1, 2, 3)$ градуса. Углы поворота на этапе случайного поиска задаются генератором случайных чисел, а на этапе функционирования — обученной нейронной сетью.

Алгоритм квазислучайного поиска прост:

- 1) начинается случайное движение рукой в произвольном направлении;
- 2) если при движении в выбранном направлении модули разностей проекций между рукой и целевой точкой на обеих сетчатках не возрастает, то рука сохраняет это направление;
- 3) если один или оба модуля разности проекций начинают возрастать, то выбирается новое случайное направление и возврат на пункт 2. Процесс повторяется до тех пор, пока «кисть» не достигнет заданной окрестности целевой точки. Описанная выше система с «рукой», «глазами» и нейросетью в дальнейшем для краткости будет называться в данном контексте аниматором

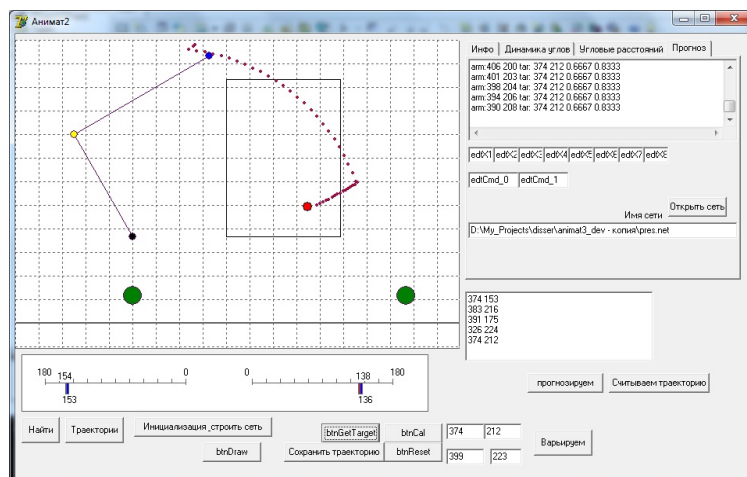


Рис. 1. Интерфейс виртуального стенда

Обучение нейронной сети

Обучение сети производилось по сохраненной потактовой записи значений входов и выходов во время случайного поиска.

Нейросеть состояла из 3 слоев и имела 9 нейронов во внутреннем слое. Обучение происходило в программном пакете FANN Tools [URL: www.leenissen.dk/fann/wp]. Количество слоев и нейронов были подобраны эмпирическим путем, метод обучения, и типы активационных функций были выбраны на основе анализируемых данных с помощью собственного функционала программы FANN Tools. Тип обучения, активационные функции и инициализационные коэффициенты сети были постоянны для всех экспериментов в процессе работе. Это было сделано для исключения влияния неконтролируемых фактов в связи с различными весовыми коэффициентами и типами обучения.

Входными данными для сети являются (таблица 1) `arm_left`, `arm_right`, `target_left`, `target_right` — координаты положения руки-манипулятора и целевой точки для левого и правого глаза соответственно, на выходе `out 1`, `out 2` — нормированные управляющие команды для двух сочленений руки-манипулятора.

Функционирующая сеть должна была выдавать команды поворота шарнирам в зависимости от положения объекта и текущего положения «кисти». Выходные значения нейронной сети показывают изменения угла для совершения оптимального действия. Если команда на выходе сети $(-2:1)$, это означает, что сегмент a нужно повернуть против часовой стрелки на 2 градуса, а сегмент b — по часовой на 1 градус. Обученная сеть способна потактово выдавать последовательность команд, приводящих руку-манипулятор к цели.

Таблица 1. Фрагмент задачника для нейросети

arm_left	arm_right	target_left	target_right	out 1	out 2
0.7300	0.4200	0.6740	0.3740	1.0000	0.3333
0.7280	0.3900	0.6740	0.3740	0.1667	0.8333
0.7240	0.4180	0.6740	0.3740	0.8333	0.8333
0.7140	0.4140	0.6740	0.3740	0.1667	0.1667
0.7240	0.4180	0.6740	0.3740	0.8333	0.6667
0.7180	0.4080	0.6740	0.3740	0.8333	0.6667
0.7120	0.4000	0.6740	0.3740	0.8333	0.6667
0.7060	0.3920	0.6740	0.3740	0.8333	0.6667

Исходные траектории, как правило, далеки от оптимальных, и одним из этапов проверки предложенной гипотезы является оценка способности нейронной сети к обобщению опыта, полученного на одной единственной траектории. Обобщение опыта должно проявиться в сглаживании траектории захвата объекта по сравнению с исходной и, как следствие, в ее укорачивании.

Кроме того, чтобы предложенная гипотеза не противоречила реальности, необходимо, чтобы формируемые поведенческие паттерны обладали робастностью — устойчивостью к возмущениям входных данных. Также необходимо выяснить, существует ли связь между конфигурацией траектории исходной траектории и качеством обучения. Из общих соображений можно ожидать, что некоторые траектории легче запечатлеваются и обобщение содержащихся в них закономерностей приводит к более робастному импринту.

Результаты и обсуждения

Сглаживание траекторий

Обученные нейронные сети демонстрируют способность к сглаживанию траекторий захвата (рис. 2, 3). Этот эффект является достаточно неожиданным, поскольку критерий, в соответствии с которым нейросеть обучалась, не направлен в явном виде на сглаживание траектории.

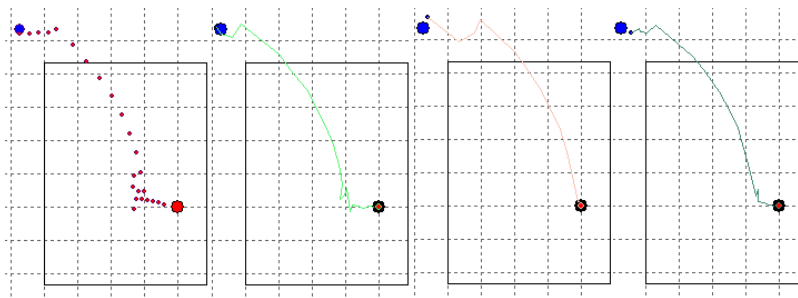


Рис. 2. Примеры сглаживания траектории захвата при варьировании начального положения «кисти». Исходная траектория (красными точками), остальные — результат работы сети. Маленькие точки в окрестности начального положения «кисти» — новые начальные положения «кисти»

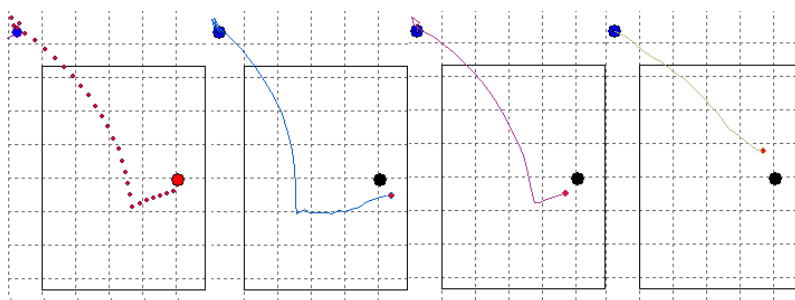


Рис. 3. Примеры сглаживания траектории захвата при варьировании положения объекта. Исходная траектория (красные точки), остальные — результат работы сети. Маленькие точки в окрестности начального положения объекта — новые положения объекта

Оценка робастности

Уже на рисунках 2 и 3 можно видеть, что анимат проявляет робастность, т. е. захват объекта осуществляется не только при точном воспроизведении начального положения «кисти» и объекта, но и при варьировании и того и другого. Количественно величину робастности можно оценить по размерам допустимых (т. е. приводящих к захвату объекта) областей начальных положений «кисти» и объекта (рис. 4).

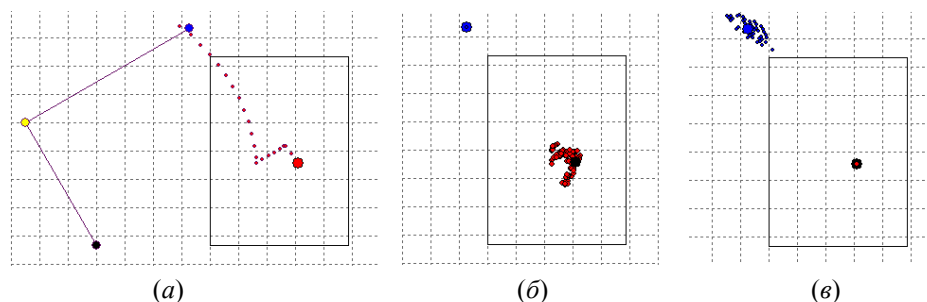


Рис. 4. Области притяжения (захвата) возникающие (а) при варьировании исходной траектории (красные точки); (б) при вариации целевой точки; (в) при вариации начального положения «кисти»

Полученные результаты подтверждают предположение о робастности модели формирования импринта, устраняя тем самым один из важных аргументов против предложенной гипотезы.

Оценка средних радиусов для разных траекторий одной целевой точки

На этом этапе работы проводилась проверка предположения, что качество обучения и устойчивость алгоритма зависят от исходной траектории. Для этого формировался набор реализаций, различающихся положением целевой точки на зрительном поле, которое задавалось генератором случайных чисел. В каждой реализации отбирались 2 траектории, конфигурации которых существенно различались (рис. 5). Затем по каждой траектории и в каждой реализации обучался соответствующий анимат. Для сравнения качества обучения был выбран средний радиус контура области захвата: чем больше средний радиус, тем выше считалось качество обучения.

Построение контуров и, соответственно, нахождение средних радиусов производились следующим образом (рис. 6): от центральной целевой точки строились лучи с углом, на 10 градусов отличающиеся от соседнего (всего 36 лучей). Вдоль каждого луча, начиная от центральной точки, через расстояние r друг от друга строились новые целевые точки, которые анимат должен был «схватить». Точки находились сетью до тех пор, пока сеть не перешагивала предел робастности, где уже была не способна захватить целевую точку. Такая операция производилась для каждого луча.

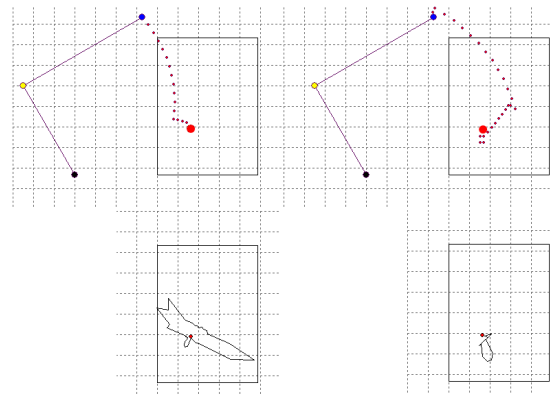


Рис. 5. Оригинальные траектории и контуры областей, в которых сети способны находить новые целевые точки ($r = 0.5$, координаты точки: (319, 183))

Интересно отметить, что область захвата не всегда обладает непрерывностью: можно видеть, что при «движении» по лучу образуются разрывы, соответствующие неспособности анимата «захватить» цель. Контур проводится соединением крайних точек лучей, выходящих из начального положения цели.

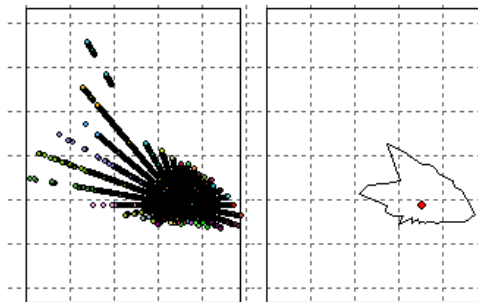


Рис. 6. Пример определения области захвата путем пошагового ($r = 0.1$) движения по лучу. Контур соединяет крайние найденные точки лучей

Для количественной оценки размеров областей притяжения траекторий без учета их формы проводилось сравнение средних радиусов отклонения [Лакин, 1990] с помощью двухвыборочного t -теста (t -критерия Стьюдента сравнения средних) с различными дисперсиями. В таблице 2 показаны сравнения средних радиусов и достоверность различия (или отсутствие достоверности).

Таблица 2. Сравнение средних радиусов для пар траекторий для уровня достоверности 5 % и $t_{кр} = 2.01$

Номер пары	Траектория № 1	Траектория № 2	Значение t -статистики
1	0.87	12.67	-7.73
2	0.42	2.77	-5.90
3	13.29	9.07	1.29
4	21.53	6.19	4.29
5	2.43	18.80	-7.54
6	17.72	23.11	-1.37
7	5.44	25.24	-7.75
8	7.22	23.23	-4.23
9	23.78	3.42	10.70
10	7.83	9.31	-1.40

Можно видеть, что в 7 экспериментах из 10 средние радиусы захвата для разных траекторий достоверно различаются. Изображения областей захвата для некоторых пар траекторий приведены на рисунке 7.

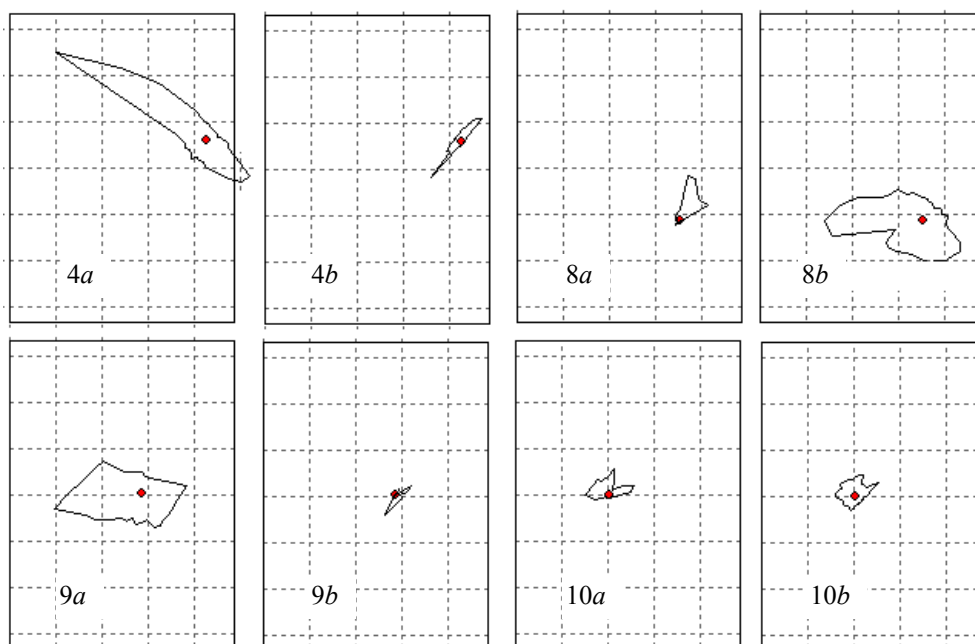


Рис. 7. Примеры областей захвата для разных траекторий и целевых точек. Изображены пары № 4, 8, 9 из таблицы 2

Поскольку топология сетей была одинакова (4-9-2) и начальные весовые коэффициенты также одинаковы, то единственным варьирующим фактором является конфигурация траекторий. Тогда остается признать, что конфигурации траекторий, достаточно часто оказывает влияние на качество функционирования сетей. Трудно удержаться от возможно поверхностной аналогии, что физическая «неуклюжесть» некоторых детей может быть следствием того, что первые исходные траектории, по которым формировались базовые паттерны, случайно оказались неблагоприятными в вышеупомянутом смысле.

Выводы

1. Предложенная модель формирования импринта продемонстрировала способность к «запечатлению» двигательного паттерна на основе предъявления всего одной траектории, приводящей к цели. Продемонстрированное моделью обобщение полученного единичного опыта, проявляющееся в сглаживании траекторий захвата, при функционировании достаточно необычно, поскольку способность супервизорных нейросетей к обобщению наблюдается при обучении на множестве примеров.

2. Оценка робастности модели продемонстрировала устойчивость к возмущению входных данных, что указывает на реалистичность предлагаемой модели.

3. Зависимость степени сглаживания траекторий захвата и их робастности (способности эффективно реагировать на более широкий набор ситуаций) от формы траекторий проявляется не всегда, но в большинстве случаев достоверно значима.

Заключение

В рамках разработанной модели удалось наблюдать свойства самопроизвольной оптимизации поведенческого паттерна на основе единичного предъявления ситуации. Эта простейшая модель позволяет наблюдать достаточно сложное поведение, в том числе обобщение исходных ситуаций, проявляющееся, помимо сглаживания траекторий движения, также в робастности отклика, т. е. способности реагировать на условия, отличающиеся от предъявленных при обучении. Данное направление исследований представляется актуальным в контексте построения реалистичных поведенческих моделей, в задачах искусственного интеллекта в рамках бионического подхода, а также в целях понимания универсальных механизмов формирования поведенческих навыков, включая импринты высших животных.

Список литературы

- 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.
- Балацкий Е. В.* Вспомогательные импринты и поведение человека // *Вестник Российской академии наук*, 2007. — Т. 77, № 10.
- Баранова А. А.* Педиатрия: национальное руководство: в 2 т. — М.: ГЭОТАР-Медиа, 2009. — Т. 1. — 1024 с.
- Гаазе-Рапопорт М. Г., Поспелов Д. А.* От амебы до робота: модели поведения. — М.: Наука. Гл. ред. физ.-мат. лит., 1987. — 288 с.
- Гаврилова Т. А., Хорошевский В. Ф.* Базы знаний интеллектуальных систем. Учебник.— СПб.: Питер, 2000.
- Гладков Л. А., Курейчик В. В., Курейчик В. М.* Генетические алгоритмы: Учебное пособие. — 2-е изд. — М.: Физматлит, 2006. — С. 320.
- Демин А. В., Витяев Е. Е.* Логическая модель адаптивной системы управления // *Нейроинформатика*, 2008. — Том 3, № 179.
- Жданов А. А.* Адаптивные машины — неизбежное направление развития техники. задачи и проблемы, научная сессия НИЯУ МИФИ–2010. XII Всероссийская научно-техническая конференция «нейроинформатика–2010»: лекции по нейроинформатике. — М.: НИЯУ МИФИ, 2010. — 328 с.
- Кривошеев С. В.* Оптимизация движения с помощью решения обратной задачи кинематики для твердотельной модели многозвенного манипулятора в среде Matlab методом перебора / С. В. Кривошеев, Д. Р. Орманов // *Автоматизация и современные технологии.* — 2012. — № 12. — С. 20–22.
- Коршунов Ю. М., Коршунов Ю. М.* Математические основы кибернетики. — М.: Энергоатомиздат, 1972.
- Лакин Г. Ф.* Биометрия: Учеб. пособие для биол. спец. вузов / 4-е изд., перераб. и доп. — М.: Высшая школа, 1990. — 352 с.
- Лоренц К.* Кантонская доктрина априори в свете современной биологии // *Журнал «Человек».* — 1997. — № 5.
- 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.
- Рассел С., Норвиг П.,* Искусственный интеллект: современный подход / 2-е изд. Пер. с англ. — М.: Издательский дом «Вильямс», 2006. — 1408 с.
- Редько В. Г.* Перспективы исследований на стыке информатики и биологии // *Нейроинформатика*, 2007. — Том 2, № 1.

- Смит К. Ю. М.* Биология сенсорных систем / Пер. с англ. — М.: БИНОМ. Лаборатория знаний, 2005. — 583 с.
- 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.
- Тарасов В. Б.* Системно-организационный подход в искусственном интеллекте // Программные продукты и системы. — 1999. — № 3. — С. 6–13.
- Хайкин С.* Нейронные сети: полный курс /2-е издание. Пер. с англ. — М.: Издательский дом «Вильямс», 2006. — 1104 с.: ил. — Парал. тит. англ.
- Carpenter G. A. and Grossberg S.* ART 3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures // *Neural Networks*. — 1990. — Vol. 3. — P. 129–152. FANN Neuro Network Library URL: www.leenissen.dk/fann/wp/ (дата обращения 13.02.2014).
- Rumelhart D. E., Hinton G. E., Williams R. J.* Learning internal representation by error propagation // *Parallel Distributed Processing*. — 1986. — Vol. 1, No. 8.
- Sutton R. S., Barto A. G.* Reinforcement learning: An introduction // Cambridge, MA: MIT Press. — 1998. — 323 p.