

Система управления автономным мобильным роботом на основе адаптивного резонанса¹

А. В. Сыцко

Аннотация. В данной работе рассматривается подход к построению автономной адаптивной системы управления с использованием искусственных нейронных сетей на примере управления мобильным роботом.

1. Введение

В настоящее время все чаще возникают задачи управления при отсутствии точных математических моделей объекта управления (ОУ) и окружающей его среды. Одним из подходов к решению этой проблемы является методология автономного адаптивного управления (ААУ), разрабатываемая в отделе имитационных систем Института системного программирования РАН [4, 5]. В данной работе основные принципы ААУ использованы для построения управляющей системы (УС) моделью автономного мобильного робота, описанной в [6]. Основным механизмом, используемым для реализации систем ААУ, являются сети из специальных нейроподобных элементов. Эти специальные модели биологических нейронов способны изменять свое функционирование в зависимости от накапливаемой ими статистики по входным векторам. Благодаря своей способности выделять неслучайные компоненты во входных данных, сети на базе таких нейронов удобны для реализации систем ААУ. Однако было бы полезно найти способы реализации систем ААУ также и на основе «классических» искусственных нейронных сетей (ИНС). Это позволило бы интегрировать метод ААУ в сферу ИНС, использовать в системах ААУ полезные свойства ИНС, привлечь специалистов по ИНС. Но, к сожалению, как выяснилось в результате данной работы, подходы, используемые в «классических» искусственных нейронных сетях, мало применимы для построения систем ААУ. Тем не менее, некоторые решения были получены. Именно, в данной работе была предпринята попытка использования традиционных нейроподобных сетей для реализации систем ААУ, а точнее одной из ее компонент УС – подсистемы формирования и распознавания образов. Для этого была использована разработанная в 1987 г. Гроссбергом и Карпенгером *теория адаптивного*

резонанса [1, 2] и ее простейшая реализация, которую в литературе принято обозначать ART-1.

2. Программная модель мобильного робота

Для исследования и демонстрации свойств УС в качестве объекта управления выбрана простейшая программная модель мобильного робота (рис. 1а), представляющего собой тележку с колесным приводом и управляемыми передними колесами, обеспечивающими движение по дуге заданного радиуса. Реверсивный движитель обеспечивает движение передним и задним ходом. За один такт времени робот может совершить передвижение по одной из 6-ти траекторий (рис. 1б). Робот имеет цилиндрический корпус и снабжен визуальными (например, лазерными или сонарными) и тактильными (бамперными) датчиками (рис 1а). Визуальные датчики обеспечивают регистрацию препятствий в одном из секторов обзора, характеризуемых углом отклонения от направления «вперед» и расстоянием от корпуса робота. Тактильные датчики регистрируют соударения модели с препятствиями.

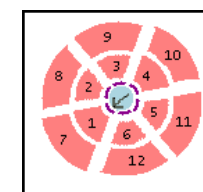


Рис. 1а.

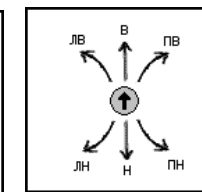


Рис. 1б.

Основная задача, ставившаяся перед УС робота – выработка стереотипов поведения при обходе стандартных препятствий, размещенных случайным способом в среде. Расположение препятствий подобрано таким образом, чтобы робот мог легко попадать в новые ситуации, но при этом чтобы соударения не происходили слишком редко. Изначально в УС не закладывается никаких априорных знаний, ни о самих препятствиях, ни о каких-либо правилах их обхода. Основываясь на «эмоциональной» оценке положения модели по отношению к окружающей среде в каждый такт, робот должен сам понять, каким показаниям каких датчиков соответствуют соударения с препятствиями и какие действия необходимо предпринять, чтобы их избежать.

3. Система управления

Управляющая система (УС) робота в данной работе была основана на принципах ААУ и с использованием известного метода reinforcement learning [8].

¹ Работа поддержана РФФИ, проекты № 03-01-00323 и № 04-01-08023-офи

Управляющая система была построена максимально независимо от конкретного ОУ и среды его обитания.

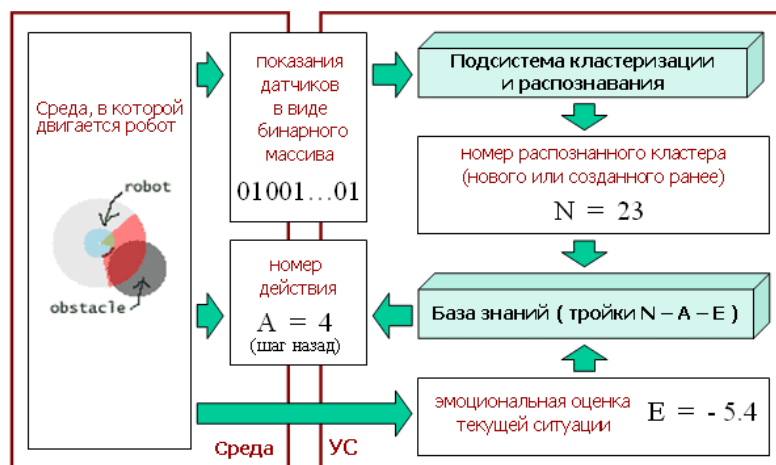


Рис. 2. Схема функционирования УС мобильного робота. В блоках показаны примеры значений переменных.

В начале каждого такта управления показания всех датчиков преобразуются в бинарную форму (например, по принципу: «1» означает, что есть препятствие в секторе, «0» - иначе) и объединяются в двоичный вектор. Положение отдельных компонент данных в этом векторе не имеет значения, так как УС сама учится понимать, что означают те или иные данные. Подобная однородность данных – одна из замечательных черт описываемого метода. Она делает алгоритм работы системы независимым от специфики используемых датчиков и, в целом, от природы входной информации.

Полученный таким способом бинарный массив поступает на вход подсистемы динамической кластеризации и распознавания. Задача этой подсистемы заключается в объединении данных в кластеры непосредственно в процессе жизни ОУ. В этом состоит принципиальное ее отличие от обычно используемых алгоритмов решения задач кластеризации, когда все возможные входные вектора предоставляются сразу. В качестве реализации этой подсистемы выбрана нейророботная сеть ART-1. Основные принципы работы ART-1 подробно описаны в [3]. Эта сеть содержит два уровня нейронов. Соответствующие друг другу пары нейронов с разных уровней, представляют собой один кластер. Веса верхнего слоя содержат нормализованный вектор – центр кластера, который модифицируется в процессе обучения. Веса нижнего слоя дискретны и являются собой «характерные черты» кластера. В процессе работы, в случае, если поступивший для кластеризации входной вектор содержит недостаточное

число *характерных черт* для того, чтобы отнести его к какому-либо из уже созданных кластеров, сеть ART-1 создает новый кластер, инициализируя его этим вектором.

На выходе подсистемы получается номер заново созданного или распознанного кластера. Этот номер можно назвать «образом». Если данный образ еще не встречался УС раньше, то УС может выбрать произвольное из возможных для ОУ действий, так как управляющая система не располагает изначально никакими правилами поведения. Если же *образ* уже встречался, то необходимо принять во внимание накопленные ранее знания о действиях ОУ в аналогичной ситуации и результатах, к которым они привели. Данная задача решается с помощью подсистемы, называемой «база знаний». База знаний осуществляет хранение и использование «знаний» в виде троек вида «образ»-«действие»-«результат». Ниже мы раскроем эти понятия.

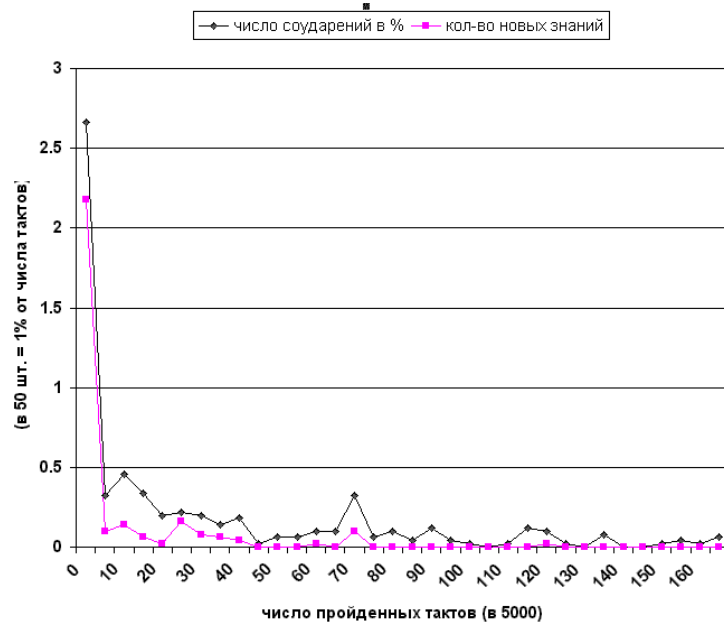
Помимо блока датчиков, ОУ взаимодействует со средой посредством актуаторов, которые УС активизирует своими выходными командами - *действиями*. В нашем случае «действие» - это одно из шести возможных движений платформы, осуществляемых колесным приводом. Как и в случае с датчиками, УС изначально не знает, к чему приводит то или иное движение, и начинает понимать это только в процессе обучения. После совершения какого-либо действия, ОУ оказывается в новой ситуации. Присутствующий в УС аппарат эмоций, основываясь на новых показаниях датчиков, выдает интегральную оценку состояния робота на текущий момент (например, если в результате предпринятого действия произошло соударение с препятствием, то это «плохо», а если не произошло, то «хорошо»). УС изначально понимает, «что такое хорошо, что такое плохо», т. е. может интерпретировать эту оценку, называемую «результатом» осуществленного действия. В этом и состоит основной принцип работы: имея в базе знаний для какого-либо распознанного образа O_1 записи о том, что действие A_1 привело к соударению, а действие A_2 не привело, УС уже обладает необходимыми знаниями для того, чтобы обойти препятствие, попавшее в поле зрения ОУ.

Таким образом, в процессе жизни ОУ решается задача максимизации эмоциональной оценки, что способствует решению глобальных задач *выживания* (в нашем случае, это избегание соударений) ОУ и *накопления знаний*.

4. Полученные результаты

Для тестирования УС была использована моделирующая система на базе модели мобильного робота «Гном № 8» [6, 7]. Ниже на графике представлено изменение числа столкновений и количества новых знаний в единицу времени (за 5000 тактов, т. е., например, число столкновений в процентах от числа тактов). Таким образом, можно видеть, что число соударений с 2,5% в начальный период жизни робота падает до 0,1 % в процессе обучения. График количества приобретенных новых знаний за те же промежутки времени повторяет график соударений. Это происходит, например, из-за

того, что число соударений может увеличиться, когда робот попадет в неисследованный им до этого участок карты, где встретится с новыми ситуациями, действовать в которых он еще не обучен, и «методом проб и ошибок» получит новые знания. Но так как препятствия расставлены в пространстве однородно, то 95% всех своих знаний робот получает в самом начале.



По нижней оси отложено число пройденных тактов в масштабе 1 : 5000. По вертикали — процент столкновений (черным) и удельное число новых приобретенных знаний (нижний график).

5. Заключение

Таким образом, продемонстрирована возможность обучения и эффективного управления ОУ, с УС сконструированной по описанным принципам. Приведенный здесь результат показывает работоспособность предложенной методологии. Свойство независимости алгоритма от семантики входных и выходных параметров должно обеспечить его применимость к решению самых различных задач. Различные УС будут отличаться только реализацией «учителя» - аппарата эмоций, который выставляет оценки действиям ОУ. Для этого, конечно, оценки должны соответствовать семантике входных данных.

Но, в то же время, показана сложность применения ИНС при решении задач ААУ. В классическом виде для использования ИНС необходима процедура обучения на заданном наборе входных векторов, что противоречит условию адаптивности. В описанной УС для реализации блока распознавания и кластеризации упомянутой УС было предпринято несколько отступлений от классической модели ART-1. Помимо этого, остальные компоненты УС, такие как, например, база знаний, были реализованы простым алгоритмическим способом. Направлением дальнейших исследований может стать поиск более универсального эффективного и изящного способа применения ИНС для реализации метода ААУ.

Литература

1. G. A. Carpenter and S. Grossberg, A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine, Computer Vision, Graphics, and Image Processing 37, 54-115 (1987)
2. S. Grossberg, Competitive Learning: From Interactive Activation to Adaptive Resonance, Cognitive science 11, 23-63 (1987)
3. Уоссерман Ф., Нейрокомпьютерная техника: теория и практика, М.: Мир (1982)
4. Жданов А. А., Арсеньев С. В., Половников В. А., Об одной методологии автономного адаптивного управления. Труды Института системного программирования РАН. 1999. Том 1. М.: Биоинформсервис, 2000.-С. 66-83.
5. Жданов А. А., Метод автономного адаптивного управления, его свойства и приложения. Интеллектуальные системы. С. 1-14
6. Жданов А. А., Крыжановский М. В., Нейронная адаптивная система управления. V всероссийская научно-техническая конференция Нейроинформатика-2003, Сборник научных трудов, Ч. 1. М.: МИФИ, 2003. С. 163-168.
7. Жданов А.А., Крыжановский М.В., Преображенский Н.Б. О возможности создания систем автономного искусственного интеллекта.// Интеллектуальные и многопроцессорные системы-2001// Тезисы докладов Международной научной конференции. Таганрог: Изд-во ТРТУ, 2001. С. 245-248. ISBN 5-8327-0089-9.
8. Richard S. Sutton, Andrew G. Barto, Reinforcement Learning: An Introduction, MIT Press, Cambridge, MA, 1998