

А.В.ГАВРИЛОВ

Новосибирский государственный технический университет
Andr_gavrilov@yahoo.com

АЛГОРИТМ ОБРАТНОГО РАСПРОСТРАНЕНИЯ ОШИБКИ ДЛЯ ОБУЧЕНИЯ НА ОТРИЦАТЕЛЬНЫХ ПРИМЕРАХ

В докладе предлагается обобщенный алгоритм обучения обратным распространением ошибки, ориентированный на обучение как на положительных, так и на отрицательных примерах. Предлагается использовать его, в частности, для решения навигационных задач мобильным роботом.

Ключевые слова: нейронные сети, обучение обратным распространением ошибки, мобильные роботы.

Введение

В последние два десятилетия в рамках робототехники наиболее бурно развивающимися направлениями являются «аниматы» (или Behavior Based Robotics [1]) и оформившееся позже направление «когнитивная робототехника» (Cognitive Robotics). В рамках этих направлений ведутся исследования и разработки в области создания роботов, обучаемых и обучающихся поведению в незнакомой среде. Очевидно, что одним из наиболее популярных и перспективных подходов к решению этой проблемы является использование нейронных сетей (краткий, но далеко не полный обзор зарубежных публикаций о применении нейронных сетей в мобильной робототехнике приведен в [2]).

Для обучения роботов на примерах наиболее удобным методом обучения является обратное распространение ошибки (Error Back Propagation - ЕВР) [3]. В работах [4, 5] предлагается использовать для управления мобильными роботами сети с прямым распространением и с алгоритмом ЕВР. Однако, этот метод имеет ряд недостатков, одним из которых, существенным для применения в робототехнике, является его ориентация на обучение с использованием положительных примеров. Это означает, что робота можно учить «как надо действовать в определенной ситуации», но нельзя учить «что не надо делать в определенной ситуации». Этим же недостатком страдает подход, основанный на использовании адаптивной резонансной теории (ART) [6] при самообучении робота, в частности, гибридная модель MLP-ART2,

предложенная автором [7], основанная на комбинации сети прямого распространения и адаптивной резонансной модели ART-2. Хотя эта гибридная модель демонстрировала некоторые преимущества [8-11] (инвариантность распознавания образов) по сравнению с использованием только MLP с ЕВР или только ART, она ориентирована также на запоминание положительных образов и больше на восприятие образов, а не на обучение поведению. В работах [12-14] также делались попытки применения модели ART для решения навигационных задач мобильным роботом. Известны и примеры применения других моделей нейронных сетей для этих целей. Например, в [15] рассмотрено использование ассоциативной памяти на основе модели Хопфилда. Недостатком ее является работа с бинарными сенсорными данными. Аналоговая сеть Хопфилда, используемая в [16], лишена этого недостатка. Однако, автору представляется более перспективным подход, основанный на нейронной сети прямого распространения (FFNN), как наиболее универсальной модели нейронной сети, для которой доказана возможность решения любых задач классификации и регрессии [17]. К тому же, эта модель легко встраивается в гибридные интеллектуальные системы управления, например, предложенную в [18].

В статье предлагается обобщенный алгоритм обучения обратным распространением ошибки, который обеспечивает обучение, как на положительных, так и на отрицательных примерах, первоначально предложенный автором в [19, 20]. Алгоритм предназначен для обучения мобильного робота поведению в незнакомой среде.

Модель робота

На рисунке 1 показана структура информационной системы для управления мобильным роботом, включающая в себя 1) нейронную сеть прямого распространения (FFNN) для выработки выходного управляющего вектора в ответ на входной вектор с дистанционных сенсоров, критик, 2) модуль-критик для определения наказания или поощрения (моделирование положительных и отрицательных эмоций), используемых для запуска одного из режимов обучения нейронной сети, 3) кратковременной памяти для хранения истории принятия решений нейронной сетью.


```

Then
    Выбор случайного направления для поворота
    Direction_to_turn;
Else
    Direction_to_turn:=WorkNN(values_from_sensors);
End if
    Сохранение входного и выходного векторов
    в кратковременной памяти;
End if
    Move(Direction_to_turn);
    Current_situation := последняя ситуация в кратковременной памяти;
    r := 1;
If Estimation имеет значение поощрение или наказание then
    While не вся кратковременная память просмотрена
        LearnNN(Current_situation, Estimation);
        r := r/2;
        Current_situation := предыдущая ситуация
        в кратковременной памяти;
    End while
End if
End while
End of Алгоритм поведения робота

```

В этом алгоритме с помощью процедуры *WorkNN* запускается нейронная сеть для получения выходного вектора, процедура *LearnNN* служит для запуска процесса обучения нейронной сети при заданной ситуации (входной и выходной векторы) и заданном режиме обучения (*Estimation*) на положительном или отрицательном примере. Процедура *Move* обеспечивает поворот робота в заданном направлении и его перемещение на один шаг.

Приведенный алгоритм обеспечивает в случае достаточно большого расстояния до препятствия принятие решения о направлении движения, используя нейронную сеть или иногда случайный выбор (для исключения возможности заикливания пути при недостаточно обученной нейронной сети). Если расстояние до препятствия мало (случай столкновения), используется простой детерминированный алгоритм отворота от него.

Обучение нейронной сети происходит в случае получения поощрения или наказания (появление в зоне прямой видимости цели или произошедшее столкновение, соответственно). При этом в качестве

примеров (положительных или отрицательных) используются входной и выходной векторы, хранящиеся в кратковременной памяти, и соответствующие предыдущим моментам принятия решения (нейронной сетью или случайным выбором). Следует иметь в виду, что обучение для каждого примера происходит с использованием сравнительно небольшого заданного количества итераций.

Обобщенный алгоритм обратного распространения ошибки

Предлагаемый обобщенный алгоритм обратного распространения ошибки (Generalized Error Back Propagation, GEBP) может работать в двух режимах - позитивном и негативном. Позитивный режим представляет собой классический ЕВР, обеспечивающий стремление к позитивным примерам-выходным векторам. Негативный режим основан на модифицированном ЕВР, обеспечивающим избегание отрицательных примеров-выходных векторов. Изменение весов в алгоритме GEBP происходят по следующей формуле:

$$\Delta w_{ij} = ar\varphi_j x'_i, \quad (1)$$

где:

w_{ij} - вес связи между i -м и j -м нейронами;

a - режим, принимает значения 1 или -1;

r - коэффициент, задающий скорость обучения;

φ_j - функция обратного распространения ошибки для j -го нейрона;

x'_i производная функции активации i -го нейрона.

Функция φ_j для вычисления ошибки, распространяемой от выходных нейронов, для случая $a=1$ соответствует классическому ЕВР:

$$\varphi_j = y_j(1 - y_j)(d_j - y_j), \quad (2)$$

где y_j и d_j - актуальный и желаемый выход выходного нейрона соответственно.

Для случая $a=-1$ функция φ_j вычисляется по формуле:

$$\varphi_j = ky_j(1 - y_j) \exp\left[-\frac{1}{2\sigma^2} (d_j - y_j)^2 \right] \quad (3)$$

Выражение $y_j(1 - y_j)$ в этой формуле представляет производную активационной функции. Экспоненциальная функция обеспечивает максимальное значение φ_j при равенстве актуального и желаемого состояния j -го нейрона. Величина σ задает чувствительность близости к избегаемому выходному вектору. Коэффициент k может интерпретироваться как уровень «страха» и может использоваться при моделировании эмоций и использовании их для обучения нейронной сети.

Другой вариант вычисления φ_j возможен следующим образом.

Для $d_j \neq y_j$ функция φ_j может быть определен как:

$$\varphi_j = \frac{ky_j(1 - y_j)}{d_j - y_j} \quad (4)$$

Для $d_j = y_j$ величина φ_j может быть равна постоянной величине k .

В отличие от классического алгоритма обучения обратным распространением ошибки обобщенный алгоритм обеспечивает обучение стремлению к «хорошему» поведению и избегание «плохого». Частным случаем может быть предвидение событий во времени и стремление к ним или избегание. В этом случае нейронная сеть прямого распространения может быть заменена на рекуррентную нейронную сеть, имеющую дело с последовательностью во времени входных и выходных векторов, например, на модель Элмана [21] с алгоритмом обучения обратным распространением ошибки во времени (EBP through time).

Эксперименты

Для проверки предложенного обобщенного алгоритма обучения методом обратного распространения ошибки и архитектуры информационной системы робота использовалось имитационное 2D-моделирование робота в среде оболочки MRS, разработанной автором с участием студентов.

При моделировании робота предполагалось, что он обладает двенадцатью дистанционными датчиками, расположенными по окружности, как показано на рис. 2а. На рис. 2б показана форма с параметрами нейронной сети (MLP) и кратковременной памяти (Memory). Моделирование осуществлялось в дискретном времени. При этом использовались следующие значения параметров:

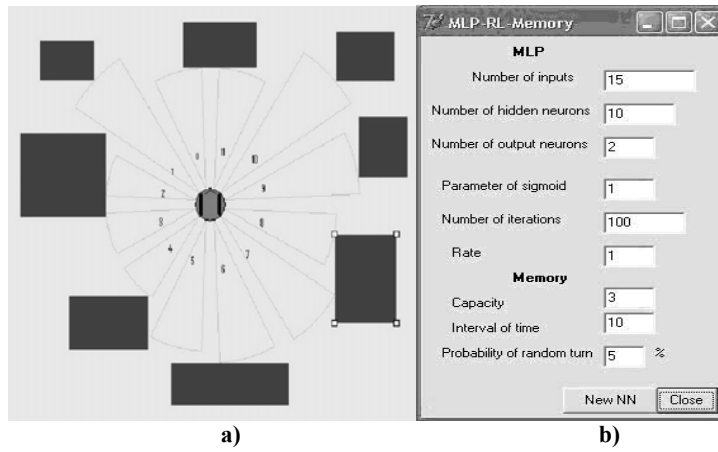


Рис. 2. а) Дистанционные сенсоры в MRS и б) параметры моделирования нейронной сети и кратковременной памяти..

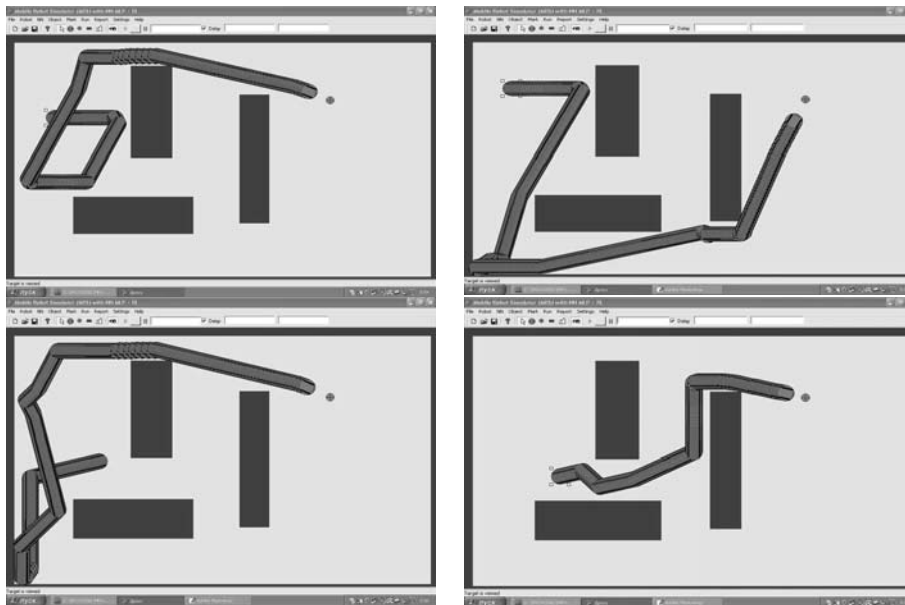


Рис. 3. Четыре сеанса моделирования робота в оболочке MRS с обучением одной и той же нейронной сети.

- 1)Интервал между моментами принятия решения - 10 тактов.
- 2)Вероятность использования случайного выбора - 0.05 (5%).
- 3)Количество итераций - 100
- 4)Количество входных нейронов - 15 (12 дистанционных сенсоров, 13-ый - направление движение, 14-ый и 15-ый – текущие координаты робота).
- 5)Количество промежуточных нейронов - 10
- 6)Количество выходных нейронов – 2 (угол поворота направо и угол поворота налево с разными знаками). Реальный поворот вычисляется как среднее арифметическое между ними.

На рисунке 3 показана серия из экспериментов с траекторией перемещения робота к цели при одной и той же нейронной сети. Следует иметь в виду, что в начале моделирования робот ориентируется на цель.

Заключение

Предложенный алгоритм обучения обратным распространением ошибки является обобщением классического алгоритма ЕВР на случай обучения на отрицательных примерах. Применительно к применению в робототехнике отрицательный пример означает принятие решения, которого надо избегать в дальнейшем.

Проведенные эксперименты показали пригодность предложенного алгоритма для самообучения (при наличии критика, формирующего поощрения и наказания) мобильного робота движению к цели с обходом препятствий.

Планируется продолжить совершенствование модели и моделирование с целью получения количественных оценок эффективности предложенного алгоритма. Также, планируется применение алгоритма GEВР на реальном роботе и применение его в предложенной автором архитектуре системы управления роботом, ориентированной на обучение с использованием естественного языка и контекста [22, 23].

Кроме того, алгоритм GEВР может быть использован для построения интеллектуальной системы, обучение которой дружественному поведению основано на использовании положительных и отрицательных эмоций [24].

Список литературы

1. Brooks R. A robust system layered control system for a mobile robot // IEEE Trans. on robotics and automation. 1986. RA-2. P. 14--23.

2. An-Min Zou, Zeng-Guang Hou, Si-Ya Fu, Min Tan. Neural Network for Mobile Robot Navigation. A Survey // ISNN-2006. Int. Symp. on Neural Networks. Proceedings. 2006. LNCS. V. 3972. Berlin, Heidelberg, New York:Springer-Verlag. P. 1218--1226.
3. Rumelhart D.E. Parallel Distributed Processing. // Explorations in the Microstructure of Cognition. Vol. I, II, McClelland J.L. (Eds.), MA, Cambridge:MIT Press. 1986.
4. Janglova D. Neural Networks in Mobile Robot Motion // Int. J. of Advanced Robotic Systems. 2007. V. 1(1). P. 15--22.
5. Golovko V., Schilling K., Roth H., Sadykhov R., Albertos P., Dimakov V. The Architecture of the Neural System for control of a Mobile Robot. // Proc. of the International Conference on Neural Networks and Artificial Intelligence. 1999. P.57 - 61.
6. Carpenter G.A. Grossberg S. Pattern Recognition by Self-Organizing Neural Networks, Cambridge. MA:MIT Press. 1991.
7. Гаврилов А.В. Гибридная модель нейронной сети на основе моделей перцептрона и ART-2 // Материалы конф. «Нейроинформатика-2005». 2005. С. 256-261.
8. Gavrilov A.V. Hybrid neural network based on models Multi-Layer perceptron and Adaptive Resonance Theory. // KORUS-2005. 9th Int. Russian-Korean Symp., Novosibirsk :NSTU. 2005. P. 604—606.
9. Gavrilov A.V., Lee Y.-K., Lee S.-Y. Hybrid Neural Network Model based on Multi-Layer Perceptron and Adaptive Resonance Theory. // ISNN06. Int. Symp. on Neural Networks. Berlin, Heidelberg, New York :Shpringer-Verlag. LNCS. Vol. 3972. P. 707—713.
10. Gavrilov A.V., Lee S.-Y. An Approach for Invariant Clustering and Recognition in Dynamic Environment. // Advances and Innovations in Systems, Computing Science and Software Engineering (Ed. Khalet Elleithy). Heidelberg :Springer. 2007. P. 47-52.
11. Gavrilov A.V., Lee S.-Y. Usage of Hybrid Neural Network Model MLP-ART for Navigation of Mobile Robot. // ICIC'07. Int. Conf. on Intelligent Computing. 2007. LNAI. V. 4682. Berlin, Heiderberg: Springer-Verlag. P. 182-191.
12. Rui Araujo. Prune-able fuzzy ART Neural Architecture for Robot Map Learning and Navigation in Dynamic environment // IEEE Trans. on Neural Networks. 2006. Vol. 17(5). P. 1235--1249.
13. Ah-Hwee Tan. FALCON: A fusion architecture for learning, cognition and navigation. // IJCNN-04. IEEE International Joint Conference on Neural Networks. 2004. Vol. 4. P. 3297-3302.
14. Сыцко А.В. Система управления автономным мобильным роботом на основе адаптивного резонанса. // Материалы XXIX Академических чтений по космонавтике. 2005. М. С. 93.
15. Billard A., Hayes G. DRAMA, a Connectionist Architecture for Control and Learning in Autonomous Robots // Adaptive Behavior. 1999. 7(1). P. 35--63.
16. Даринцев О.В., Мигранов А.Б. Использование нейронной карты для планирования траектории мобильного робота. // Искусственный интеллект. 2009. №3. С. 300-307.

17. Горбань А.Н. Обобщенная аппроксимационная теорема и точное представление многочленов от нескольких переменных суперпозициями многочленов от одного переменного // Изв. вузов. Математика. 1998. № 5. С. 6-9.
18. Гаврилов А.В., Губарев В.В., Джо К.-Х., Ли Х.Х. Архитектура гибридной системы управления мобильного робота. // Мехатроника, автоматизация, управление. 2004. №8. С. 30-37.
19. Gavrilov A., Lee S.-Y. Unsupervised hybrid learning model (UHLM) as combination of supervised and supervised models. IEEE Int. Conf. SMC UK&RI, Dublin. 2007.
20. Gavrilov A.V., Lenskiy A. Mobile Robot Navigation Using Reinforcement Learning Based on Neural Network with Short Term Memory // ICIC'11. Int. Conf. on Intelligent Computing. 2011. LNCS 6838. Berlin, Heidelberg:Springer-Verlag. P. 210-217.
21. Elman J.L. Distributed representations, simple recurrent networks, and grammatical structure. // Machine Learning. 1991. 7(2/3). P. 195-226.
22. Gavrilov A. V. Context and Learning based Approach to Programming of Intelligent Equipment. // ISDA-08. The 8th Int. Conf. on Intelligent Systems Design and Applications. 2008. Taiwan. P. 578-582..
23. Gavrilov A.V. New Paradigm of Context based Programming-Learning of Intelligent Agent. // Proc. of 1st Workshop on Networked embedded and control system technologies. In conjunction with 6th Int. Conf. on Informatics in Control, Automation and Robotics ICINCO-2009. 2009. Portugal:INSTICC PRESS. P. 94-99.
24. Gavrilov A.V. Emotions and a priori Knowledge Representation in Artificial General Intelligence. INFOS-2008. Int. Conf. on Intelligent Information and Engineering Systems Varna. 2008; In "Intelligent Technologies and Applications" of Int. Book Series "Information Science and Computing", Bulgaria :ITHEA. P. 106-110.