

## АЛГОРИТМ ОБУЧЕНИЯ МНОГОСЛОЙНОГО ПЕРСЕПТРОНА НА ОТРИЦАТЕЛЬНЫХ ПРИМЕРАХ

А.В.ГАВРИЛОВ\*, А.А. ЯКОВЕНКО♥

В данной работе предлагается алгоритм обучения перцептрона методом «проб и ошибок» с наказанием в случае неправильной реакции на входной вектор. Этот алгоритм может использоваться в тех случаях, когда нельзя определить «правильное» поведение, но можно определить «неправильное» поведение. Такой случай может быть при управлении мобильным роботом или аниматом, функционирующим в незнакомой среде.

**1. ВВЕДЕНИЕ.** В последние 10-15 лет получили распространение исследования в области так называемого обучения с подкреплением или методом «проб и ошибок» (reinforcement learning) [1], ориентированного на применение в системах управления мобильных роботов или аниматов [2] и при создании «искусственной жизни». При использовании обучения с подкреплением интеллектуальная система учится правильно себя вести в окружающей среде. Обычно для его реализации используются нейронные сети [3-4]. Из часто используемого русскоязычного названия этого подхода к обучению («с подкреплением») видна его направленность на обучение на основе положительных примеров, т.е. при его использовании система учится «как надо поступать» в той или иной ситуации. В основе такого обучения как правило лежит использование перцептрона и какой-либо разновидности алгоритма обучения обратным распространением ошибки (error back propagation) или ассоциативной памяти на основе модели, подобной модели Хопфилда [5].

В то же время для организации поведения мобильного робота или анимата часто необходимо обеспечить обучение «на отрицательных примерах», т.е. учиться «как не надо поступать» в определенных ситуациях, например, при решении задачи избегания (убегания) [6].

---

\* Доцент кафедры вычислительной техники, канд. техн. наук, докторант  
♥ магистрант

В статье предлагается алгоритм обучения методом «проб и ошибок» для многослойного персептрона с прямыми связями, веса связей которого представляют собой вещественные числа в диапазоне от -1 до 1. При этом результатом работы такого персептрона считается не вектор вещественных чисел, полученных на его выходах в ответ на некоторый входной вектор, а выход с максимальным значением сигнала. В качестве примера применения такой сети можно привести задачу выбора стратегии поведения в зависимости от сложившейся ситуации, где каждой стратегии соответствует свой выход нейронной сети, а ситуация представляется в виде вектора вещественных чисел (признаков), подаваемого на входы нейронной сети.

**2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ АЛГОРИТМА.** Принцип работы предлагаемого алгоритма обучения многослойного персептрона заключается в следующем: после получения входного вектора нейронной сети и вычисления его выходного вектора выбирается выходной нейрон с максимальным значением выхода (этот выбор и считается откликом сети). Если отклик можно считать верным в данной ситуации, то обучения (изменения весов связей) не происходит. В противном случае сети посылается «сигнал боли», запускающий процесс корректировки весов связей между нейронами в направлении от выходов к входам в зависимости от предыдущих значений весов, уровня сигналов, приходящих на вход каждого нейрона и порядкового номера данного слоя по отношению к выходному слою сети. Веса «ошибочно» сработавших нейронов уменьшаются. Веса всех остальных выходных нейронов (сработавшие каждого из которых теоретически может оказаться «правильным» в данной ситуации) увеличиваются. Величина изменения веса каждого входа каждого нейрона прямо пропорциональна уровню сигнала, пришедшего на данный вход, предыдущему значению веса данного входа и среднему значению произведения величины сигнала на выходе данного нейрона на веса соответствующих входов всех нейронов следующего слоя. Кроме того, данная величина обратно пропорциональна удаленности данного слоя от выходного слоя сети (своеобразное «затухание» «импульса боли»).

Разработанный алгоритм предназначен для обучения классического многослойного персептрона с прямыми связями, поведение которого описывается следующими зависимостями:

$$y_j = \frac{s_j}{|s_j| + a}; \quad s_j = \sum_i x_i w_{ij}(t);$$

где:  $y_j$  – значение сигнала на выходе нейрона;  $w_{ij}(t)$  – значение весового коэффициента для  $i$ -го входа  $j$ -го нейрона в момент времени  $t$ ;  $x_i$  – уровень сигнала, пришедшего на  $i$ -й вход ( $0 < x_i < 1$ );  $a$  – константа ( $0 < a < 1$ ); в проведенных экспериментах  $a = 0.5$ .

Новое значение весового коэффициента  $i$ -го входа  $j$ -го нейрона в момент времени  $t$  при обучении персептрона вычисляется как

$$w_{ij}(t) = w_{ij}(t-1) + r x_i w_{ij}(t-1) g_j,$$

где:  $w_{ij}(t)$  – новое значение веса  $i$ -го входа  $j$ -го нейрона;  $w_{ij}(t-1)$  – предыдущее значение того же весового коэффициента;  $x_i$  – уровень сигнала, пришедшего на  $i$ -й вход ( $0 < x_i < 1$ );  $r$  – коэффициент скорости обучения ( $0 < r < 1$ );  $g_j$  – коэффициент поощрения/наказания для  $j$ -го нейрона, который вычисляется следующим образом:

а) для выходного слоя:  $g_j = -1$  для «неверно» сработавшего нейрона (т.е. обладающего максимальным уровнем сигнала на выходе среди остальных нейронов выходного слоя для данной ситуации) и  $g_j = 1/N$  для остальных нейронов (здесь  $N$  – количество нейронов в выходном слое);

б) для всех остальных слоев:  $g_j = y_j \sum_k w_{jk}(t) g_k$ , где суммирование

ведется по всем нейронам следующего слоя, связанным по входам с  $j$ -м нейроном.

**3. ЭКСПЕРИМЕНТЫ.** Предложенный алгоритм был реализован и проверен в виде двух программ:

- модель сражения мобильного робота с несколькими другими;
- модель ассоциативной памяти, формируемой при обработке простых графических изображений.

В обоих случаях использовался персептрон с одним скрытым слоем.

В первой модели робот подвергается преследованию нескольких роботов-противников. При этом робот может выбирать один из двух видов поведения – отход (в каком-либо направлении) и нападение (на одного из противников). Результатом нападения робота на робота-противника является уменьшение «жизненной силы» последнего. В качестве входного вектора используются координаты противников в окрестности робота. Сигнал наказания вырабатывается в случае уменьшения «жизненной

силы» робота. Эксперименты показали, что робот достаточно хорошо обучается выбору такого поведения, что пытается нападать, когда это возможно, на одиночного противника и убегать, если есть опасность нападения сразу нескольких противников.

Во второй модели алгоритм обрабатывает серию клеточных графических образов с заданием для каждого из них множества возможных «правильных» выходов (выбранных выходных нейронов). Сигнал наказания вырабатывается в случае несовпадения полученного отклика с одним из возможных «правильных».

**4. ЗАКЛЮЧЕНИЕ.** Предварительные результаты экспериментов показали, что при использовании предложенного алгоритма нейронная сеть достаточно хорошо обучается на отрицательных примерах. В дальнейшем планируется продолжить исследования в следующих направлениях:

- попытаться увеличить эффективность алгоритма,
- разработать математическое обоснование предлагаемого алгоритма и оценки его эффективности,
- апробировать предложенный алгоритм при построении компьютерной игры и на программной и реальной моделях мобильного робота.

[1] DIETTERICH T.G., FLANN N. *Explanation-based learning and reinforcement learning: A unified view* - // Proc. of the 20th Int. Conf. on Machine Learning, Morgan Kaufmann, California, 1995 – Pp. 176-184.

[2] НЕПОМНЯЩИХ В.А. *Поиск общих принципов адаптивного поведения живых организмов и аниматов.* - Новости искусственного интеллекта, № 2(50), 2002. - С.48-53.

[3] HALLAM B. *Animal Learning Models as Robot Controllers.* – // Proc. of the SIRS2000, MIT Press, 2000. – Pp. 131-140.

[4] 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, Brest, 1999. - Pp.57 - 61.

[5] КРУГЛОВ В.В., БОРИСОВ В.В. *Искусственные нейронные сети. Теория и практика.* – М.: Горячая линия-Телеком, 2001.

[6] GAVRILOV A.V. *An Architecture of Neurocomputer for Image Recognition.* - Neural Network World, N.1,1991. – Pp. 59-60.

