

Модель обучения в иерархической структуре функциональных систем

П.Л. Отоцкий¹, В.Г. Редько²

¹Московский физико-технический институт (государственный университет)

²Институт оптико-нейронных технологий РАН

peter@dgap.mipt.ru

Анимат-подход к разработкам искусственного интеллекта сложился сравнительно недавно, в начале 1990-х годов. Эта методология исследует особенности формирования интеллекта с помощью моделирования искусственных организмов (аниматов) в переменных внешних средах. Одна из главных проблем исследований аниматов – постараться понять процессы формирования целенаправленного адаптивного поведения, исследовать процессы возникновения нетривиальной структуры целей и подцелей в интеллектуальном поведении. Мы планируем провести такие исследования с точки зрения теории функциональных систем, развитой в 1930-1970-х годах советским нейрофизиологом П.К. Анохиным [1]. А именно, мы строим модель в рамках проекта «Мозг Анимата» [2], который базируется на теории функциональных систем и рассматривает иерархические схемы обучения.

Мы рассматриваем поведение анимата в модельной двумерной клеточной среде (рис. 1), которая использовалась для анализа процессов эволюционного возникновения иерархической структуры целей и подцелей в работе [3].

Мы предполагаем, что система управления аниматом имеет иерархическую архитектуру (рис. 2). Базовым элементом системы управления является отдельная функциональная система (ФС).

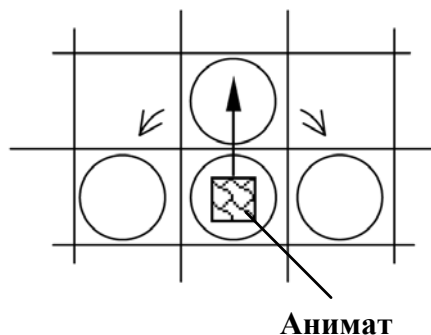


Рис. 1. Анимат в двумерной клеточной среде. Анимат ориентирован (стрелка показывает направление вперед), кружки – поле зрения анимата. Действия анимата: двигаться вперед, поворачиваться направо или налево, есть, бороться с другими аниматами.

Предполагаем, что система управления аниматом функционирует в дискретном времени и каждый такт времени активна только одна ФС. Мы рассматриваем простую формализацию функциональных систем на основе нейросетевых адаптивных критиков [4]. Адаптивные критики – достаточно простые схемы управления, обучаемые с помощью хорошо известного метода обучения с подкреплением [5]. Формализованная ФС моделирует следующие важные особенности ее биологического прототипа: 1) прогноз результата действия, 2) сравнение прогноза и результата, и 3) коррекцию прогноза путем обучения в соответствующих нейронных сетях.



Рис. 2. Архитектура системы управления аниматором. Каждый блок представляет собой функциональную систему, реализованную на основе конструкции адаптивного критика и предназначенную для выбора одного из возможных действий.

Каждая ФС включает в себя адаптивный критик, который состоит из двух блоков: Модель и Критик. Предполагаем, что Модель и Критик – многослойные нейронные сети, и что производные по весам нейронных сетей могут быть вычислены обычным методом обратного распространения ошибки. Мы также предполагаем, что в каждый момент времени адаптивный критик должен выбрать одно из возможных действий с учетом оценки качества прогнозируемых ситуаций. Роль Модели – прогноз следующей ситуации для всех возможных действий. Роль Критика – оценка качества ситуации $V(S)$ для текущей ситуации $S(t)$, следующей ситуации $S(t+1)$ и прогнозируемых ситуаций $S^{pr}_i(t+1)$ для всех возможных действий. Детальное описание работы используемой схемы адаптивных критиков содержится в [2].

В настоящее время реализуется программа модели рассматриваемой схемы обучения.
Работа поддержана РФФИ, проект 04-01-00179.

Литература:

1. П.К. Анохин. Принципиальные вопросы общей теории функциональных систем // Принципы системной организации функций. – М.: Наука, 1973.
2. V.G. Red'ko, D.V. Prokhorov, M.S. Burtsev. Theory of functional systems, adaptive critics and neural networks // International Joint Conference on Neural Networks, Proceedings. Budapest, 2004. PP. 1787-1792.
3. М.С. Бурцев. Модель эволюционного возникновения целенаправленного адаптивного поведения. 2. Исследование развития иерархии целей // Препринт ИПМ РАН, 2002, № 69.
4. В.Г. Редько, Д.В. Прохоров. Нейросетевые адаптивные критики // Научная сессия МИФИ-2004. VI Всероссийская научно-техническая конференция "Нейроинформатика-2004". Сборник научных трудов. Часть 2. М.: МИФИ, 2004. С.77-84.
5. R. Sutton, A. Barto. Reinforcement Learning: An Introduction. – Cambridge: MIT Press, 1998.