

КЛАССИЧЕСКАЯ И ЭВОЛЮЦИОННАЯ ПРИЧИННОСТЬ В МОДЕЛЯХ ОБУЧЕНИЯ*

М.С.Бурцев¹

В работе рассматривается проблема причинности в моделях обучения. Предложено выделять два типа детерминации обучения – классический и эволюционный. Проведена формализация и сравнение этих типов причинности. Рассмотрены проблемы построения моделей обучения основанных на эволюционной причинности, намечены возможные пути их решения.

Введение

Наиболее популярные методы обучения, используемые в моделях поведения животных [Balkenius C. et al., 1999; Dayan P. et al., 2002; Joel D. et al., 2002; Spier E. et al., 1997; Suri R.E. et al., 2001] и при создании искусственных программных и робототехнических адаптивных систем [Blumberg B., 1996; Sutton R.S. et al., 1998], основываются на предположении, что обучение однозначно детерминируется внешними или внутренними причинами. Такое понимание природы причин, приводящих к изменению поведения, аналогично пониманию причинности в классической физике, назовем этот тип причинности классическим или детерминистским. Подобный тип причинности не является единственно возможным [Мамчур Е.А. и др., 2002; Campbell D.T. et al., 2003; Поппер К.Р., 2000а; Пригожин И., 1999], примерами современных научных теорий, в которых используются другие виды причинности, могут служить квантовая механика и эволюционная теория.

Применение классической причинности при исследовании поведения животных критиковалось, как со стороны нейрофизиологии [Анохин П.К., 1935; Сахаров Д.А., 2004], так и со стороны концептуальных подходов к проблемам самоорганизации и живого [Поппер К.Р., 2002; Пригожин И., 1999] и обучения [Поппер К.Р.; 2000b]. Возможно ли моделирование обучения основанного на причинности отличной от причинности используемой в большинстве современных моделей? Если да, то, как

* Работа выполнена при финансовой поддержке РФФИ (проект № 04-01-00179)

¹ 125047, Москва, Миусская пл., д.4, ИПМ им. М.В. Келдыша РАН, mbur@narod.ru

нужно строить такую модель, и какие проблемы при этом возникнут? Ниже будет приведена попытка формализовать два возможных различных типа причинности в обучении и сравнить их.

Формализация причинности в обучении

Для формализации процесса обучения используем автоматный подход [Цетлин М.Л., 1969]. Рассмотрим агента, имеющего внутреннюю среду и действующего во внешней среде. Представим "мозг", управляющий поведением агента, в виде автомата $S\{S_E, S_I, S_S, S_A, F\}$ (см. рис.1). Здесь S_E – множество элементов автомата, состояния которых определяются внешней по отношению к агенту средой; S_I – элементы, определяющиеся внутренней средой агента; S_S – внутренние элементы автомата; S_A – состояния, определяющие действия агента; F – функция при помощи, которой вычисляются состояния элементов автомата. В каждый момент времени автомат описывается вектором состояний всех его элементов S_k

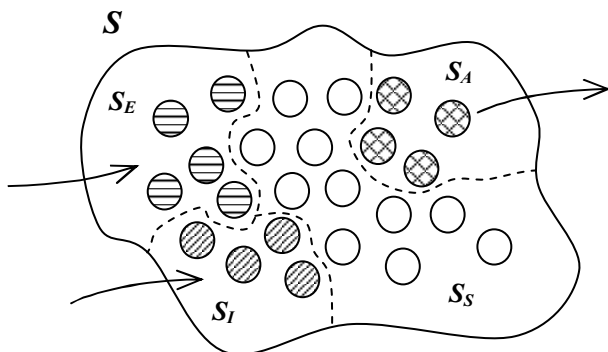


Рис. 1. Схема автомата, управляющего поведением агента. S_E – множество элементов автомата, состояния которых определяются внешней по отношению к агенту средой; S_I – элементы, определяющиеся внутренней средой агента; S_S – внутренние элементы автомата; S_A – элементы, определяющие действия агента

Будем считать, что наш автомат функционирует в дискретном времени $t=(t_0, t_1, t_2, \dots, t_n)$. Подобное допущение принимается в большинстве моделей адаптивного поведения. Тогда поведение некоторого автомата S_k можно представить как:

$$\dots \longrightarrow S_k^{t_n} \longrightarrow S_k^{t_{n+1}} \longrightarrow S_k^{t_{n+2}} \longrightarrow \dots, \quad (1)$$

или:

$$S_k^{t_{n+1}} = F_k(S_E^{t_{n+1}}, S_I^{t_{n+1}}, S_S^{t_n}, S_A^{t_n}). \quad (2)$$

Таким образом, мы определили поведение через функцию, преобразующую состояния автомата S . В общем случае эта функция

зависит от состояния внешней и внутренней среды агента в текущий момент времени, а так же от внутренних состояний самого автомата S_S и S_A в предыдущий момент времени. Если функция F_k некоторого автомата S_k не изменяется от одного момента времени к другому, то обучения не происходит. Назовем обучением процесс изменения функции F . Одновременное поведение и обучение рассматриваемого автомата можно записать следующим образом:

$$\dots \longrightarrow S_k^{t_n} \xrightarrow{T(F_k)} S_k^{t_{n+1}} \xrightarrow{T(F_k)} S_k^{t_{n+2}} \xrightarrow{T(F_k)} \dots, \quad (3)$$

или:

$$F_{k^{(n+1)}} = T(F_{k^{(n)}}), \quad (4)$$

здесь T – некоторое преобразование, действующее на множестве возможных функций F . Добавив в описание автомата преобразование T , получаем автомат с обучением $S\{S_E, S_I, S_S, S_A, F, T\}$.

В общем виде преобразование T может действовать не только на F , но и на множество состояний автомата. Дополним (4) и запишем соотношения, определяющие поведение и обучение автомата:

$$\left. \begin{aligned} S_{k^{(n+1)}} &= T(S_{Ek^{(n)}}, S_{Ik^{(n)}}, S_{Sk^{(n)}}, S_{Ak^{(n)}}, F_{k^{(n)}}) \\ S_{k^{(n+1)}}^{t_{n+1}} &= F_{k^{(n+1)}}(S_{Ek^{(n+1)}}^{t_{n+1}}, S_{Ik^{(n+1)}}^{t_{n+1}}, S_{Sk^{(n)}}^{t_n}, S_{Ak^{(n)}}^{t_n}) \end{aligned} \right\}. \quad (5)$$

В рамках приведенного формализма обучение определяется через задание вида функционала T , назовем его функционалом или алгоритмом обучения. Основная масса методов обучения, применяемых в искусственных нейронных сетях, а также алгоритмы обучения с подкреплением, не изменяют множество состояний автомата, т.е. преобразование T в этих моделях обучения действует только на F (см. формулу (4)).

Обучение и классическая причинность

Основное предположение, определяющее вид функционала T , в популярных методах обучения, заключается в том, что процесс обучения детерминирован собственными состояниями автомата. В большинстве алгоритмов T зависит от внутренних и внешних входов автомата, хотя в некоторых работах допускается зависимость от внутренних состояний автомата [Ziemke T., 2000]. Следовательно, доминирующая парадигма основана на зависимости следующего вида:

$$T = f(S_E^{t_n}, S_I^{t_n}, S_S^{t_n}, S_A^{t_n}), \quad (6)$$

т.е. изменение поведения (функции определяющей поведение) (4) является функцией от состояний самого автомата.

Обучение и эволюционная причинность

Какую альтернативу можно предложить классической причинности в обучении? Основное допущение альтернативного подхода к обучению заключается в следующем утверждении. Алгоритм обучения *не зависит* от состояний самого автомата. Попробуем построить такой алгоритм.

Для того чтобы T не являлось функцией от $\mathcal{S}$, представим его следующим образом:

$$T(\mathcal{S}) = \{S_E, S_I, S_S, S_A, F\} \otimes \xi, \quad (7)$$

здесь ξ – случайный процесс (шум), а $\otimes$ – обозначает воздействие на множество состояний автомата и функцию поведения F . Другими словами обучение заключается в случайных модификациях поведения автомата $\mathcal{S}$.

Теперь процессы поведения и обучения автомата $\mathcal{S}_k$ можно представить так:

$$\mathcal{S}'_k \xrightarrow{S_k \otimes \xi} \mathcal{S}'_k \xrightarrow{S_k \otimes \xi} \dots \xrightarrow{\xi \otimes \xi} \xi \xrightarrow{\xi \otimes \xi} \dots \quad (8)$$

Очевидно, что если мы воспользуемся преобразованием (7) в общем виде, то автомат будет эволюционировать к такому виду, в котором каждое действие будет выбираться случайно с равной вероятностью. Следовательно, необходимо так изменить (7), чтобы при сохранении независимости самого преобразования от состояний автомата, не происходило деградации поведения. Заметим, что преобразование (7) является формальным описанием порождения разнообразия, если мы будем рассматривать процесс эволюции автоматов $\mathcal{S}$. Шум ξ служит источником "мутаций" при копировании автомата. Получается, что процесс функционирования автомата (8) выглядит, как постоянное "мутирование" автомата, но в эволюционных процессах кроме мутаций, также присутствует отбор.

Будем считать, что отбор в процессе обучения реализуется за счет изменения амплитуды шума-"мутаций" ξ . Тогда (7) можно записать, как:

$$T(\mathcal{S}) = \{S_E, S_I, S_S, S_A, F\} \otimes A\xi, \quad (9)$$

где A – величина амплитуды шума. Причинность, задаваемая выражением (9) заключается в создании вариаций и отборе, назовем такой вид детерминации изменений эволюционным. Естественно, что отбор при обучении должен чем-то управляться, а значит, амплитуда шума должна от чего-то зависеть. Сделаем еще одно допущение конкретизирующее (7) и (9).

Предположим, что отбор управляется самим агентом, а точнее состояниями его "мозга" $\mathcal{S}$. Выразим это формально:

$$T(\mathcal{S}) = \{S_E, S_I, S_S, S_A, F\} \otimes A(S_s)\xi. \quad (10)$$

Выражение (10) значит, что агент самостоятельно решает, когда учиться, а когда нет. При этом после всех преобразований, выражение (10) соответствует основному тезису эволюционной причинности, что T не зависит ни от состояния внешней и внутренней среды агента, ни от состояния его "мозга". Внутренние состояния автомата S могут управлять лишь величиной преобразований при обучении, но не направленностью самого обучения. Другими словами, агент в каждый момент времени решает, стоит ли модифицировать свое поведение или нет, и если решено изменить поведение, то на основе уже имеющегося опыта генерируется новая *случайная* проба.

Два противоположных подхода (6) и (10) к детерминации обучения обладают принципиально различными свойствами. Если в первом случае, обучение зависит от уже имеющегося опыта, направляется и ограничивается им, что приводит к "оптимизации" поведения в рамках уже известного агенту. В этом случае отсутствует возможность приобретения "новых" знаний. Напротив, при эволюционной детерминации обучения всегда присутствует случайность, которая служит источником новизны [Чернавский Д.С., 2001], обеспечивает "открытость" обучения, отсутствие границ процесса приобретения новых знаний.

Открытые проблемы

Хотя выражение (10) задает направление поиска альтернативных алгоритмов обучения основанных на эволюционной причинности, но для реализации конкретного метода обучения в виде модели его явно не достаточно. Для построения модели необходимо ответить, по крайней мере, на два вопроса.

1. Как происходит воздействие "шума" на "мозг"? Или, как операция $\otimes$ действует на $\{S_E, S_I, S_S, S_A, F\}$?
2. Как агент должен решать для себя, в каких ситуациях ему нужно учиться, а в каких нет? Или, каков вид зависимости $A(S_S)$?

Готовых ответов на эти вопросы сегодня не существует. Попытаемся наметить пути поиска возможных решений.

Обучение будет эффективным, если уже имеющиеся знания не разрушаются, а используются для создания новых решений. В качестве алгоритма, соответствующего этим условиям, можно предложить следующий:

1. вставка новых элементов (внутренних состояний) в автомат;
2. установление функциональной зависимости (расширение F) между вставленными элементами и элементами, имеющими отношение к текущему поведению (при использовании в качестве автомата искусственной нейронной сети это, например, могут быть нейроны с наибольшей активацией).

Такой алгоритм должен приводить в модели к явлению похожему на специализацию нейронов в мозге животных. Также этот алгоритм вставки может быть проинтерпретирован в терминах теории функциональных систем [Анохин П.К., 1973]. Можно сказать, что этот алгоритм описывает системогенез, в процессе которого, происходит создание новых пробных интеграций (функциональных систем) за счет добавления новых элементов, выступающих в роли акцептора результата действия.

Для поиска ответа на вопрос о том, когда стоит учиться, а когда нет, сделаем допущение, что обучение происходит только в том случае, когда уже имеющихся знаний не хватает для генерации необходимого поведения. (для искусственных адаптивных агентов, цель функционирования которых задается разработчиком, разработчик и задает функцию изменяющую амплитуду шума в зависимости от успешности поведения (целевая функция)) В теории функциональных систем такая ситуация соответствует невозможности достижения необходимого результата при помощи имеющегося репертуара функциональных систем. Однако, механизмы подобной оценки невозможности достижения результата мозгом животных практически не исследованы (что выступает в качестве дополнительного стимула к попыткам их моделирования). Для поиска $A(S_s)$ можно попытаться использовать эволюционное моделирование в надежде на то, что в процессе искусственной эволюции будет найдено решение функционально гомологичное существующему в природе. Для этого в автомате необходимо ввести специальный выход, который будет управлять амплитудой стохастического поиска. Функциональная зависимость этого выхода от внутренних состояний автомата будет находиться в процессе эволюции обучающихся агентов.

Основные особенности популярных алгоритмов и предложенной альтернативы сведены в таблицу 1.

Предложенный подход к моделированию обучения обладает рядом существенных отличий от наиболее распространенных моделей обучения. Это делает актуальной задачу его реализации в виде компьютерной модели, и последующего сравнения его динамики с динамикой обучения реальных животных и других моделей (в том числе с динамикой моделей, лежащих в области синтеза теории функциональных систем и обучения с подкреплением [Red'ko V.G. et al., 2004]). Успехи методов эволюционного синтеза искусственных нейронных сетей [Nolfi S. et al., 2002; Ruppini E., 2002], которые концептуально близки к предложенному методу обучения, позволяют надеяться, что будут получены интересные результаты.

	популярные алгоритмы обучения	обучение на основе эволюционной причинности
<i>детерминация обучения</i>	обучение зависит от состояний автомата (входов от внешней и внутренней среды, внутренних состояний)	обучение не зависит от состояний автомата, а происходит под воздействием случайной величины, амплитуда которой регулируется самим агентом
<i>деградация памяти</i>	накопленные знания изменяются	накопленные знания дополняются
<i>элементы, вовлекаемые в обучение</i>	активные в данный момент элементы	то же
<i>моменты времени, в которые происходит обучение</i>	постоянно (все моменты времени)	когда результат не может быть достигнут имеющимися средствами.

Список литературы

- [Dayan, P. et al., 2002] Dayan P. and Balleine B.W. Reward, motivation, and reinforcement learning // Neuron 36[2]. 2002.
- [Suri, R. E. et al., 2001] Suri R.E., Bargas J., and Arbib M.A. Modeling functions of striatal dopamine modulation in learning and planning // Neuroscience 103[1]. 2001.
- [Spier E. et al., 1997] Spier E. and McFarland D. Possibly optimal decision-making under self-sufficiency and autonomy // J.Theor.Biol. 189[3]. 1997.
- [Balkenius C. et al., 1999] Balkenius C. and Moren J. Dynamics of a classical conditioning model // Autonomous Robots 7. 1999.
- [Joel D. et al., 2002] Joel D., Niv Y., and Ruppini E. Actor-critic models of the basal ganglia: new anatomical and computational perspectives // Neural Netw. 15[4-6]. 2002.
- [Sutton R.S. et al., 1998] Sutton R.S. and Barto A.G. Reinforcement learning an introduction. Cambridge, Mass, MIT Press, 1998.
- [Blumberg B., 1996] Blumberg B. Old Tricks, New Dogs: Ethology and Interactive Creatures. Ph.D. MIT Media Lab, 1996.

- [Пригожин И., 1999] Пригожин И. Конец определенности. Время, хаос и новые законы природы. Ижевск, Ижевская республиканская типография, 1999.
- [Поппер К.Р., 2000] Поппер К.Р. Мир предрасположенностей: два новых взгляда на причинность. // Эволюционная эпистемология и логика социальных наук: Карл Поппер и его критики. Под ред. Садовского В.Н., Эдиториал УРСС, 2000.
- [Мамчур Е.А. и др., 2002] Причинность и телеономизм в современной естественно-научной парадигме // под ред. Мамчура Е.А. и Сачкова Ю.В. М., Наука, 2002.
- [Campbell D.T. et al., 2003] Campbell D.T. and Bickhard M.H. Variations in Variation and Selection: The Ubiquity of the Variation-and-Selective-Retention Ratchet in Emergent Organizational Complexity // Foundations of Science 8[3]. 2003.
- [Сахаров Д.А., 2004] Сахаров Д.А. Уроки малых сетей. Лекция на семинаре "Мозг", Москва 23-03-2004. См. <http://www.neurogene.ru/>.
- [Анохин П.К., 1935] Анохин П.К. Проблема центра и периферии в современной физиологии нервной деятельности // Проблема центра и периферии в нервной деятельности. Горький, 1935.
- [Поппер К.Р., 2002] Поппер К.Р. Об облаках и часах // Объективное знание. Эволюционный подход. М., Эдиториал УРСС, 2002.
- [Поппер К.Р., 2000] Поппер К.Р. Эволюционная эпистемология // Эволюционная эпистемология и логика социальных наук: Карл Поппер и его критики. Под ред. Садовского В.Н., Эдиториал УРСС, 2000.
- [Цетлин М.Л., 1969] Цетлин М.Л. Исследования по теории автоматов и моделированию биологических систем. М., Наука, 1969.
- [Ziemke T., 2000] Ziemke T. On 'Parts' and 'Wholes' of Adaptive Behavior: Functional Modularity and Diachronic Structure in Recurrent Neural Robot Controllers // From animals to animats 6 - Proceedings of the Sixth International Conference on the Simulation of Adaptive Behavior. Cambridge, MA, MIT Press, 2000.
- [Чернавский Д.С., 2001] Чернавский Д.С. Синергетика и информация: Динамическая теория информации. М., Наука, 2001.
- [Анохин П.К., 1973] Анохин П.К. Принципиальные вопросы общей теории функциональных систем // Принципы системной организации функций. М., Наука, 1973.
- [Red'ko V.G. et al, 2004] Red'ko V.G., Prokhorov D.V., Burtsev M.S. Theory of Functional Systems, Adaptive Critics and Neural Networks // Proceedings of the International Joint Conference on Neural Networks (IJCNN '04), Hungary, Budapest. 2004.

- [Ruppin E., 2002]** Ruppin E. Evolutionary autonomous agents: a neuroscience perspective // Nat.Rev.Neurosci. 3[2], 2002.
- [Nolfi S. et al., 2002]** Nolfi, S. and Floreano, D. Synthesis of autonomous robots through evolution // Trends Cogn Sci. 6[1]. 2002.