



Artificial societies. 2013-2020

ISSN 2077-5180

URL - <http://artsoc.jes.su>

All right reserved

Issue 4 Volume 14. 2019

Development of a computer habit-forming algorithm

V. Istratov

*Central Economics and Mathematics Institute
Russian Federation, Moscow*

Abstract

Despite the long history of studying the topic of habit and the interest it gets from several areas of knowledge, the process of habit formation remains underexplored. Modeling of this process by mathematical and computer simulation methods brings about particularly many difficulties. This article seeks to find the most adequate way to formalize habit formation. Giving preference to computer simulation, the author discusses the requirements for a computer algorithm of habit formation and proposes an algorithm based on qualitative categories.

Keywords list (en): habit, algorithm, simulation, behavior, decision making

Date of publication: 19.12.2019

Acknowledgment:

Работа поддержана грантом Российского фонда фундаментальных исследований (проект 18-010-01091).

Citation link:

Istratov V. Development of a computer habit-forming algorithm // Artificial societies. 2019. V. 14. Issue 4 [Electronic resource]. Access for registered users. URL: <https://artsoc.jes.su/s207751800007579-5-1/> (circulation date: 28.09.2020). DOI: 10.18254/S207751800007579-5

1 Исследователи давно обратили свое внимание на человеческие привычки и отметили их важность. В доказательство часто приводят цитаты из Уильяма Джеймса, например, о том, что «на девяносто девять сотых или, возможно, на девятьсот девяносто девять тысячных наша деятельность совершенно автоматическая и осуществляемая по привычке, начиная с того, как

мы встаем утром, и заканчивая тем, как ложимся каждую ночь» [6]. Хотя некоторых размахистость его высказываний не столько убеждает, сколько настораживает.

2 Существует множество определений привычки. В большинстве случаев они отличаются нюансами, которые для данной работы не принципиальны. Поэтому мы будем отталкиваться от достаточно общего понимания привычки как автоматизированного действия, выполнение которого в определенных условиях стало потребностью¹.

3 Исследователи обратили свое внимание на привычки не менее трех веков назад, но до сих пор не существует единой общепринятой теории ее формирования и развития. На сегодняшний день наибольший интерес к изучению привычек проявляют несколько научных областей: психология, физиология и медицина. Кроме того, привычки часто упоминаются в экономической литературе [2], но их отдельная проработка в экономических источниках почти не встречается. Нет и полной ясности у физиологов, которые продолжают поиски участков мозга, ответственных за формирование привычки. Основные претенденты – это области полосатого тела и коры, а главный кандидат на роль ключевого нейромедиатора – дофамин (например, [4]). Что касается непосредственно механизма формирования привычки (например, демонстрирующего, как вещества перемещаются из клетки в клетку и возникает привычка, или какого-то иного), то вариантов предлагается очень мало [9].

4 Хотя в том, что касается роли привычки, ученые на сегодня более-менее единодушны: привычки позволяют экономить вычислительные ресурсы мозга при решении повторяющихся задач. При этом какую именно роль привычки играют в поведении, не до конца понятно. Традиционно существовал полярный подход к принятию решений: выполнение действия может быть обусловлено либо привычкой, либо сознательной ориентацией на определенную цель [5]. Но в последнее время активно развиваются взгляды, согласно которым осуществление действия определяется как привычкой, так и стремлением к цели (например, [7]).

5 К сожалению, очень редко встречается описание процесса формирования привычки, удобного для представления его алгоритмически или математически [3]. Более того, с точки зрения точности прогнозов (а это один из важнейших критериев качества любой модели), ни одна из известных нам моделей не способна надежно предсказывать возникновение привычки. Имеется в виду, что несовпадения с реальностью по времени и условиям формирования привычки могут быть настолько большими, что лишат такой прогноз смысла и надежд на практическое применение. Поэтому в большинстве случаев прогнозы даже не делаются.

6 Несмотря на дефицит гипотез полных механизмов, можно выделить важные факторы и признаки привычки, которые необходимо учесть при разработке алгоритма ее формирования. Прежде всего, это повторяемость деятельности. Привычка формируется путем многократного выполнения одного и того же действия в тех же самых или очень схожих обстоятельствах. Вопрос, насколько схожи должны быть эти обстоятельства, остается на сегодняшний день открытым, потому что в естественных условиях формирование привычки наблюдать чрезвычайно сложно, а лабораторные эксперименты почти всегда проводят при неизменных условиях.

7 Кроме того, для построения реалистичной модели важно понимать, какие обстоятельства из всех, с которыми сталкивается человек, можно считать ключевыми для формирования той или иной привычки. И если в лаборатории все условия регламентированы экспериментатором и предположительно маловажные факторы исключаются, то в реальных условиях выбрать из всех возможных элементов среды, поведения и восприятия значимые – задача огромной сложности, не решенная в общем виде.

8 Тем не менее, главный вопрос применительно к повторяемости – это сколько нужно повторов, чтобы сформировалась привычка? На самом деле такая интуитивно возникающая формулировка является некорректной и вводит в заблуждение, поскольку число повторов далеко не единственный и, возможно, даже не главный фактор. Ибо очевидно, что у одних людей одни и те же привычки формируются быстрее (грубо говоря, после меньшего числа повторов), чем у других. Причем разница во времени формирования может измеряться годами. Однако существуют трудности с выделением других значимых факторов. Резонно предположить, что свою роль играют физиологические особенности и предрасположенности (скажем, некоторые привычки вообще не способны сформироваться у определенных людей), разнесенность актов выполнения действия во времени (чем реже выполняется действие, тем медленнее формируется привычка), степень заинтересованности в результате действия, которая к тому же может меняться со временем (чем желательнее результат, тем быстрее формируется привычка), психологическое состояние (скажем, стресс или усталость замедляют формирование привычки). Но пока исследований поддерживающих или опровергающих эти предположения не достаточно.

9 Второй важный фактор формирования привычки, который хотелось бы выделить, – это результативность выполненного действия. В более ранних подходах ему не уделялось должного внимания. Однако на сегодня очевидно, что степень успешности действия играет очень важную роль в формировании привычки. И тут возникает проблема: что называть успешным действием? В лабораторных экспериментах эта проблема не встречается, поскольку ее исключают на уровне дизайна эксперимента: допустимые действия предельно упрощаются, лишаясь неоднозначности. Например, действие «нажатие красной кнопки» может приводить к предоставлению пищи. В таком эксперименте предельно ясно, что есть действие и что есть его успех и неуспех. В реальности результаты действий, критерии их успешности и даже сами действия редко так же четко определены. Как преодолеть эту неоднозначность, не очевидно (и это на фоне заочной дискуссии о значимости успеха для формирования привычки в принципе).

10 Учет результата, как правило, происходит в моделях обучения с подкреплением. Выделяются два подхода: на основе модельного и безмодельного управления. В случае модельного управления считается, что человек строит в голове когнитивную карту (т.е. модель, обычно представляемую при помощи графа) возможных исходов и находит по ней оптимальный путь к своей цели. Таким образом имитируется целеориентированное поведение. Напротив, при безмодельном управлении человек не строит в голове никаких карт, а научается благодаря подкреплению (как правило, положительному) реагировать некоторым образом на определенные внешние стимулы. Тем самым моделируется привычное поведение.

11 Кроме того, важными признаками привычки называют два свойства: то, что исполнение привычного действия не реагирует на уменьшение значимости результата, и то, что исполнение привычного действия не реагирует на изменение результата (например, [8]). В первом случае речь идет о том, что человек, у которого сформировалась привычка, будет продолжать выполнять привычное действие, даже если результат этого действия потеряет для него ценность. Во втором случае речь идет о том, что человек повторяет привычное действие, даже когда оно перестает приводить к желаемому результату. Во многом по этим признакам современные психологи отделяют целеориентированную деятельность от привычной.

12 Считается, что подход, основанный на подкреплении, успешно воспроизводит первое свойство и плохо справляется с отражением второго [8]. Таким образом, мы получаем минимальный набор требований к алгоритму формирования привычки. Во-первых, привычка должна возникать путем повторения действия. Во-вторых, действие должно повторяться при одних и тех же обстоятельствах – в одинаковых условиях. В-третьих, успешность действия

должна играть важную роль в оформлении привычки. В-четвертых, сформированная привычка должна быть малочувствительна к уменьшению значимости результата привычного действия. В-пятых, сформированная привычка должна быть малочувствительна к разрушению связи привычного действия и результата. Первые три требования касаются особенностей методологии, последние два – содержания результатов, но вместе они позволяют глубже проработать модель привычки. Это не значит, что данный набор является достаточным, но, по крайней мере, он учитывает общие аспекты во взглядах разных исследователей.

13 Вообще говоря, в задачах компьютерного моделирования неявно присутствует еще одно требование: поскольку алгоритм привычки не может существовать в вычислительном вакууме и должен быть частью какой-то другой модели (как правило, модели, описывающей поведение), то он должен быть совместим с этой моделью. Совместим концептуально, а не формально (используя лишь единые названия, единый пользовательский интерфейс и т.п.), т.е. основываться на общих теоретических допущениях и концепциях.

14 Если говорить об общем подходе к способу представления привычки, то, на наш взгляд, предпочтение стоит отдавать методам явного представления (т.е. таким, которые описывают привычку отдельным параметром, переменной или классом) из-за их наглядности. Хотя неявное представление (например, когда привычка определяется, как возникновение определенных динамических эффектов) по форме кажется ближе к тому, как привычка кодируется в нашем мозге, наглядность позволяет эффективно верифицировать и отлаживать алгоритм и недвусмысленно интерпретировать. Поэтому на ранних исследовательских этапах наглядность приобретает дополнительную ценность.

15 Нет полной ясности и в том, что касается воплощения алгоритма формирования привычки. Прежде всего, на наш взгляд, стратегический выбор между сугубо математическим и программным моделированием следует делать в пользу последнего: в поведенческих проблемах не обойтись лишь математическим описанием, зато в программе, где необходимо, можно использовать математические формулы. Другими словами, сомнительно, что можно полноценно представить проблему формирования привычек в одной из традиционных математических форм (системы уравнений, оптимизационной задачи, теоретико-игровой задачи и т.д.). Необходима алгоритмизация, т.е. добавление элементов программного алгоритма.

16 Если говорить только о программных реализациях, то среди исследователей нет единого мнения, какому конкретному алгоритму или даже общему алгоритмическому подходу отдать предпочтение [3]. Кроме того, надо помнить, что привычки – это элемент системы принятия решений. Поэтому большая ответственность ложится и на алгоритм принятия решений, используемый в той программе, для которой разрабатывается алгоритм привычки. Друг без друга они малоценны: проявления привычки возможны только в рамках системы принятия и осуществления решений. Таким образом, то, как будут удовлетворены четвертое и пятое требования, зависит не только от самого алгоритма формирования привычки.

17 На наш взгляд, в отличие от того же Джеймса, у сознания очень много дискретных аспектов. И потому моделировать его при помощи непрерывных (а уж тем более гладких, дифференцируемых) функций – это слишком сильное допущение. Ясно, что это допущение делается вынужденно, от беспомощности. Когда не ясно, с какой стороны браться за проблему, инстинктивно хочется вооружиться проверенным в других областях знаний инструментом, а аппарат дифференцируемых функций хорошо разработан и предлагает очень удобные вычислительные возможности. И несмотря на то, что по сравнению с дискретными непрерывные функции зачастую предполагают гораздо более громоздкие вычисления, тем не менее это типовые вычисления, зачастую не требующие даже минимальной изобретательности.

18 Поэтому, на наш взгляд, было бы оправдано попробовать применить один из дискретных подходов, например, использующий качественные понятия [1]. Так, в свете этого подхода формирование привычки будет состоять из нескольких этапов. Начало первого этапа формирования привычки наступает с начала работы алгоритма. Этот этап охватывает период «ознакомления» с действием, поэтому этап будет считаться завершенным, когда окажется, что действие выполняется «легко» выполнимо. Кроме того, должны выполняться общие условия, неизменные на протяжении всех этапов:

- 19 • результат действия должен быть желателен тому, кто его выполняет (если результат безразличен или нежелателен, то привычка не формируется);
- действие должно давать ожидаемый результат (если действие совершается, но не приносит желаемого результата, то привычка не формируется);
- ключевые обстоятельства выполнения действия должны быть неизменны.

20 Далее мы предполагаем, что привычка зарождается в момент, когда хорошо знакомое действие (и его результат) переводится в разряд менее важных и мозг концентрируется на другой деятельности, но при этом все обстоятельства выполнения знакомого действия сохраняются.

21 Второй этап будет считаться завершенным, когда успешно выполнится еще одно условие: действие впервые исполнено «бессознательно». На практике это может быть реализовано через понижение важности привычного действия относительно прочих действий (т.е. мы неявно исходим из того, что сознание концентрируется только на совершении самого важного действия), но при этом запускается процесс его исполнения, которое не должно встретить препятствий, требующих участия сознания (например, вмешательства третьих лиц).

22 Третий этап будет считаться завершенным, когда все условия второго этапа будут выполнены X раз подряд. После завершения третьего этапа будем считать, что привычка сформирована. Понятие «легкого» выполнения действия, использованного при описании первого этапа, можно доопределить в рамках той же качественной концепции. Действие становится легким тоже в три этапа:

- 23 1. Первый этап пройден, когда действие впервые выполнено (т.е. проверено, что оно выполнимо).
2. Второй этап пройден, когда действие впервые выполнено без малейших шероховатостей (не возникло никаких проблем, требующих вовлечения сознания в их решение).
3. Третий этап пройден, когда условия второго этапа выполнены X раз подряд.

24 По завершении третьего этапа будем считать, что действие выполняется «легко».

25 В обоих случаях X обозначает гипотетическое минимально необходимое число повторов для возникновения нужного свойства. Несложно заметить, что при изложенном подходе удовлетворяются все пять требований к алгоритму формирования привычки – все они прописаны в явном виде.

Remarks:

1. Мещеряков Б.Г., Зинченко В.П. (ред.) Большой психологический словарь. Изд. 4-е. М.: АСТ, Прайм-Еврознак, 2009.

References:

1. Istratov V.A. Strukturno-tselevoj podkhod k kachestvennym otsenkam // Ehkonomika i matematicheskie metody. 2018, T.54, № 2.
2. Istratov V.A. Kontseptsii privychki v ehkonomicheskoy teorii i ikh prigodnost' dlya algoritmizatsii // Zhurnal Novoj ehkonomicheskoy assotsiatsii. 2019, T. 41, № 1.
3. Istratov V.A. Formalizatsiya opisaniya privychki: obzor podkhodov // Ehkonomika i matematicheskie metody. V pechati.
4. Daw N.D., Niv Y., Dayan P. Uncertainty-Based Competition between Prefrontal and Dorsolateral Striatum Systems for Behavioral Control // Nature Neuroscience. 2005, Vol. 8, No. 12.
5. Dolan R.J., Dayan P. Goals and Habits in the Brain // Neuron. 2013, Vol. 80
6. James W. Talks to Teachers on Psychology and to Students on Some of Life's Ideals // In W.James. "Writings, 1878-1899". – New York: Literary classics of the United States. 1992 P. 705-888.
7. Malvaez M., Wassum K.M. Regulation of habit formation in the dorsal striatum // Current Opinion in Behavioral Sciences. 2018, Vol. 20 P. 67–74.
8. Pauli W.M., Cockburn J., Pool E.R., Perez O.D., O'Doherty J.P. Computational Approaches to Habits in a Model-free world // Current Opinion in Behavioral Sciences. 2018, Vol. 20, P. 104–109.
9. Read S.J., Brown A.D., Wang P., Miller L.C. The Virtual Personalities Neural Network Model: Neurobiological Underpinnings // Personality Neuroscience. 2018, Vol. 1, E. 10.

Разработка компьютерного алгоритма формирования привычки

Истратов В. А.

ЦЭМИ РАН

Российская Федерация, Москва

Аннотация

Несмотря на продолжительную историю изучения темы привычек и интерес к ней со стороны сразу нескольких областей знаний, процесс формирования привычки по-прежнему остается недоисследованным. Особенно много сложностей возникает с моделированием этого процесса математическими и компьютерными методами. Данная статья посвящена поиску наиболее адекватного способа формализации процесса формирования привычки. Отдавая предпочтение компьютерному моделированию, автор обсуждает требования, предъявляемые к компьютерному алгоритму формирования привычки, и предлагает собственный вариант алгоритма, основанный на использовании качественных категорий.

Ключевые слова: привычка, алгоритм, компьютерное моделирование, поведение, принятие решений

Дата публикации: 19.12.2019

Ссылка для цитирования:

Истратов В. А. Разработка компьютерного алгоритма формирования привычки // Искусственные общества. 2019. Т. 14. Выпуск 4 [Электронный ресурс]. Доступ для зарегистрированных пользователей. URL: <https://artsoc.jes.su/s207751800007579-5-1/> (дата обращения: 28.09.2020). DOI: 10.18254/S207751800007579-5