

Российская академия наук  
Институт прикладной физики  
Центр оптико-нейронных технологий НИИСИ  
Межрегиональная ассоциация когнитивных исследований  
Нижегородский государственный университет им. Н. И. Лобачевского  
Нижегородская государственная медицинская академия  
ЗАО «Нижегородское агентство наукоемких технологий»

# **НЕЛИНЕЙНАЯ ДИНАМИКА В КОГНИТИВНЫХ ИССЛЕДОВАНИЯХ – 2013**

**ТРУДЫ  
III ВСЕРОССИЙСКОЙ КОНФЕРЕНЦИИ**

Нижний Новгород  
ИПФ РАН  
2013

## Программный комитет конференции

- Крыжановский Борис Владимирович** (Центр оптико-нейронных технологий НИИСИ РАН, Москва) – *председатель*
- Александров Юрий Иосифович** (Институт психологии РАН, Москва) – *зам. председателя*
- Анохин Константин Владимирович** (ГУ НИИ нормальной физиологии им. П.К. Анохина РАМН, Курчатовский НБИК-центр, Москва) – *зам. председателя*
- Сергеев Александр Михайлович** (ИПФ РАН, Нижний Новгород) – *зам. председателя*
- Черниговская Татьяна Владимировна** (Санкт-Петербургский государственный университет, Санкт-Петербург; Курчатовский НБИК-центр, Москва) – *зам. председателя*
- Яхно Владимир Григорьевич** (ИПФ РАН, Нижний Новгород) – *зам. председателя*
- Антонец Владимир Александрович** (ИПФ РАН, Нижний Новгород)
- Величковский Борис Митрофанович** (Курчатовский НБИК-центр, Москва)
- Гурбатов Сергей Николаевич** (ННГУ, Нижний Новгород)
- Григорьева Вера Наумовна** (НГМА, Нижний Новгород)
- Дорожкин Александр Михайлович** (ННГУ, Нижний Новгород)
- Дунин-Барковский Виталий Львович** (Центр оптико-нейронных технологий НИИСИ РАН, Москва)
- Жданов Александр Аркадьевич** (ИТМиВТ РАН, Москва)
- Еськов Валерий Матвеевич** (Сургутский государственный университет, Сургут)
- Иваницкий Алексей Михайлович** (ИВНДиНФ РАН, Москва)
- Казанцев Виктор Борисович** (ИПФ РАН, Нижний Новгород)
- Кибрик Андрей Александрович** (МГУ им. М.В. Ломоносова, Москва)
- Крылов Андрей Константинович** (Институт психологии РАН, Москва)
- Хилько Александр Иванович** (ИПФ РАН, Нижний Новгород)
- Подладчикова Любовь Николаевна** (НИИ НК ЮФУ, Ростов-на-Дону)
- Полевая Софья Александровна** (ННГУ, НГМА, Нижний Новгород)
- Парин Сергей Борисович** (ННГУ, Нижний Новгород)
- Ратушняк Александр Савельевич** (КТИ ВТ СО РАН, Новосибирск)
- Редько Владимир Георгиевич** (Центр оптико-нейронных технологий НИИСИ РАН, Москва)
- Станкевич Лев Александрович** (СПИИ РАН, Санкт-Петербург)
- Самарин Анатолий Иванович** (НИИ НК ЮФУ, Ростов-на-Дону)
- Чернавский Дмитрий Сергеевич** (ФИ РАН, Москва)
- Чернавская Ольга Дмитриевна** (ФИ РАН, Москва)
- Шахов Борис Евгеньевич** (НГМА, Нижний Новгород)
- Щербаков Виталий Иванович** (Педагогический университет, Нижний Новгород)

### Редакционная коллегия:

*В.А. Антонец, Н.Н. Кралина, И.В. Нуйдель, С.Б. Парин,  
С.А. Полевая, А.М. Сергеев, В.Г. Яхно*

Конференция организована при поддержке Российского фонда фундаментальных исследований (грант № 13-02-06119).

Электронный адрес оргкомитета: [nuidel@appl.sci-nnov.ru](mailto:nuidel@appl.sci-nnov.ru)  
Веб-сайт конференции: <http://www.nd-cogsci.iapras.ru/2013/>

# Логика работы мозга и причинность

Е.Е. Витяев

Институт математики им. С.Л. Соболева СО РАН, Новосибирск  
Новосибирский государственный университет  
vityaev@math.nsc.ru

## 1. Причинность

Принцип опережающего отражения действительности, сформулированный П.К. Анохиным, подразумевает причинность наблюдаемого мира. Что понимается под причинностью? Причинность является следствием *физического детерминизма*: «для всякой изолированной физической системы произвольно фиксированное состояние системы детерминируют все последующие состояния» [1]. Но возьмем, например, автомобильную аварию [2], что явилось ее причиной? Это может быть состояние поверхности дороги, ее влажность, расположение солнца относительно взоров водителей, нарушение правил дорожного движения, психологическое состояние водителей, неисправность тормозов и т. д. Понятно, что в этом случае нет определенной причины происшествия. Понятие причинности анализировал Д. Юм, но, как правильно отмечается в [3], он «не нашел никакого другого основания, кроме некоторого внутреннего чувства привычки».

В философии науки причинность сводится к предсказанию и объяснению. «Причинное отношение означает предсказуемость... в том смысле, что, если полная предыдущая ситуация будет известна, событие может быть предсказано... если будут даны все относящиеся к событию факты и законы природы» [2]. Понятно, что все факты, число которых потенциально бесконечно, и всех законов никто знать не может. Некоторые из законов могут быть обнаружены путем анализа данных. Поэтому причинность сводится к предсказанию в соответствии с индуктивно-номологическим выводом, основанном на выводе предсказаний из фактов и вероятностных законов.

При обнаружении законов (закономерностей) на реальных данных возникает проблема статистической двусмысленности, которая состоит в том, что в процессе обучения (индуктивного вывода) мы можем получать вероятностные правила, из которых выводится противоречие. Чтобы избежать противоречий, Гемпель [4] ввел требование максимальной специфичности законов, состоящее, образно говоря, в том, что закон должен учитывать максимум информации, относящейся к предсказываемому свойству.

Для обнаружения максимально специфических правил и избавления от противоречий нами разработан специальный семантический вероятностный вывод (ССВ) [5–7]. Доказано [5], что если в индуктивно-номологическом выводе делать предсказания, используя только максимально специфические законы и факты, то противоречий не возникнет. Таким образом, решается проблема

статистической двусмысленности и получается формализация причинности.

Для реализации принципа опережающего отражения действительности мозг в первую очередь должен уметь обнаруживать причинные зависимости. С нашей точки зрения смысл деятельности нейронов как раз и состоит в обнаружении причинных связей. В работе [8] приведена формальная модель нейрона, реализующая правило Хебба в виде семантического вероятностного вывода, а также способная обнаруживать причинные зависимости в виде максимально специфических правил.

## 2. Понятие цели и теория функциональных систем

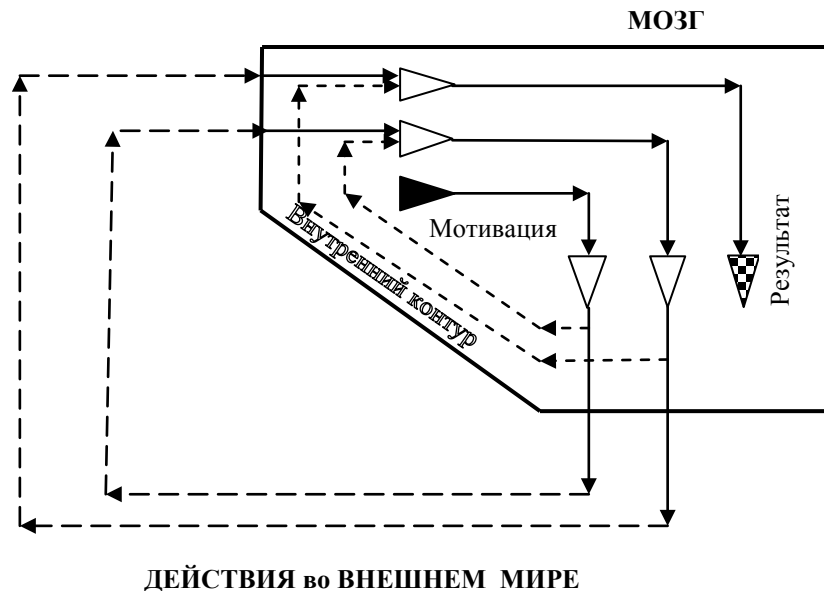
Рассмотрим целенаправленное поведение. Оно возникает, когда организмы начинают двигаться. Растения не двигаются и получают только то, что им дается. Если организм может двигаться, то он стремится изменить поступающую к нему стимуляцию, пищу и энергию нужным для него образом. Поэтому *любое действие с необходимостью становится целенаправленным* – оно стремится изменить приходящую к организму стимуляцию, что и задает цель действия. Цель нельзя достичь, не имея критерия ее достижения, иначе всегда можно считать, что цель уже достигнута и действовать не нужно. Поэтому с необходимостью должен существовать критерий достижения цели, являющийся критерием остановки действия.

Понятие цели парадоксально – цель ничего не говорит о том, *как* ее достичь и *как* организовать целенаправленное поведение. Для того чтобы знать, как достичь цели, нужен опыт. Если нет никаких знаний и опыта, то поведение организуется методом проб и ошибок. Опыт складывается из полученных методом проб и ошибок случаев достижения цели.

При целенаправленных действиях предсказания должны определять, что, как и когда можно достичь в целенаправленном поведении. Понятие цели и результата является центральным в теории функциональных систем (ТФС), где анализируется физиологический механизм цели, целеполагания и целенаправленной деятельности. Рассмотрим взаимосвязь причинности и ТФС.

**Нейрофизиологическое обеспечение причинности в ТФС.** Нейрофизиологически предвосхищение реализуется специальными коллатеральными ответвлениями от произведенных действий, которые поступают на «вход» мозга, конвергируя с афферентацией от входных стимулов [9, с. 97].

Фактически это означает выработку условных (причинных) связей между осуществлением дей-



Формирование акцептора результатов действия

ствий (эффекторным возбуждением) и последующим восприятием результатов действий, представленных их афферентными признаками (см. рисунок). Осуществляя действия, мы сразу же по коллатералям посылаем условный сигнал о том, что сейчас получим афферентацию о результатах этих действий. Это приводит к выработке условных (причинных) связей между действиями и их результатами, отражающих связи действий и результатов, происходящих во внешнем мире. Эти условные связи, осуществляемые мозгом по внутреннему контуру (см. рисунок), позволяют прогнозировать результаты действий, происходящих во внешнем мире, еще до появления самих результатов. Когда мотивационным возбуждением активируются различные последовательности действий по достижению поставленной цели по «внутреннему контуру», то выработанными условными связями прогнозируется и вся последовательность и иерархия результатов, которые могут быть получены в процессе достижения цели. Когда принято решение об определенном плане действий, то одновременно по «внутреннему контуру» предвосхищается достижение всех промежуточных результатов, которые составляют акцептор результатов действия. Можно показать, что эти причинные связи лежат в основе модели функциональных систем [10].

### 3. Принцип сенсорных коррекций Н.А. Бернштейна

В работе [11, с. 25–26] Н.А. Бернштейн пишет о том, что при большом числе степеней свободы практически невозможно рассчитать движение сложной кинематической системы, например, руки: «...при многих степенях свободы у системы суммируются... погрешности, приносимые каждой из степеней свободы; при большом количестве последних суммарная ошибка сможет вырасти до такой величины, которая покроет все преимуще-

ства, в принципе создаваемые богатым разнообразием подвижности сложной цепи. Например, если каждая из степеней свободы руки и пальца пианиста, сидящего за инструментом, даст погрешность всего в  $1^\circ$ , то, суммируясь, эти погрешности смогут дать отклонение кончика пальца на 5–6 см (хотя по отдельным звеньям, например, пальцевых фаланг, составляющие погрешности не превысят при этом 0,05 см), т. е. вызовут промахивание на терцию или кварту... Еще более существенное значение имеют осложнения динамические. В сложной кинематической цепи, каждое звено которой обладает известной тяжелой и инертной массой, всякая сила, возникающая в одном из звеньев, тотчас же вызывает целую систему реактивных или отраженных сил, передающихся на все остальные звенья. Это взаимное влияние звеньев цепи друг на друга во всех мыслимых сочетаниях создает в общей совокупности огромное количество силовых взаимодействий, совершенно необозримое математически и представляющее непреодолимые трудности для аналитического решения. Эти реактивные силы наслаиваются на те силы, которые находятся в распоряжении организма для управления движениями системы, и на внешние силы, подвластные ему всегда лишь в большей или меньшей степени, и делают общую динамическую картину движения цепи чрезвычайно осложненной, а главное – практически непредсказуемой из-за их крайней механической запутанности».

Как мозг справляется с этой проблемой? «Решение вопроса о неоднозначности лежит в использовании для регулирования эффе́кторного процесса сенсорных сигналов о позе кинематической цепи и о мере растяжения каждой из влияющих на ее движения мышц. ... Действительно, как только орган, находящийся под действием внешних и реактивных сил, плюс еще какая-то добавка внутренних мышечных сил отклонится в своем результирующем движении от того, что входит в

намерения центральной нервной системы, эта последняя получит исчерпывающую сигнализацию об этом отклонении, достаточную для того, чтобы внести в эффекторный процесс соответственные адекватные поправки. Весь изложенный принцип координирования заслуживает поэтому названия *принципа сенсорных коррекций*» [11, с. 34].

Из принципа сенсорных коррекций с необходимостью следует важный вывод: планировать определенную последовательность действий заранее невозможно, планировать и предсказывать можно только последовательность достигаемых результатов, а выбор того или иного действия должен осуществляться в реальном режиме времени по принципу сенсорных коррекций на основании поступающей обратной афферентации от уже осуществленных действий. Поэтому конкретный план действий должен фиксировать только последовательность и иерархию результатов, которые надо получить для достижения конечной цели.

В работах [10, 12–13], продолжая проект «мозг анимата», разработана формальная модель упомянутых информационных процессов и проведены эксперименты, подтверждающие ее эффективность. Эта модель следующим образом учитывает сказанное:

1) использует формальную модель нейрона, обнаруживающую причинные связи и основанную на семантическом вероятностном выводе;

2) осуществляет постановку цели в целенаправленном поведении, формирует функциональную систему и акцептор результатов действия;

3) автоматически формирует подцели и подкрепляет достижение подцелей, если их достижение увеличивает вероятность достижения конечной цели;

4) моделирует рефлекторное кольцо и сенсорные коррекции;

5) выбирает действие в реальном режиме времени с учетом текущей ситуации и получаемой афферентации;

6) сравнивает различные способы достижения цели;

7) принимает решение об определенном способе достижения цели и всех подцелей в соответствии с последовательностью и иерархией функциональных систем.

Эта работа поддержана Российским фондом фундаментальных исследований (грант № 11-07-00560-а), интеграционными проектами СО РАН № 3,

87, 136 и программой президента Российской Федерации поддержки научных школ НШ-276.2012.1.

## Литература

1. Закон. Необходимость. Вероятность. М. : Прогресс» 1967. С. 366.
2. Карнап Р. Философские основания физики. М. : Прогресс, 1971. С. 388.
3. Редько В.Г. Эволюция, нейронные сети, интеллект. Модели и концепции эволюционной кибернетики. М. : ЛИБРОКОМ, 2011. С. 220.
4. Hempel C.G. Maximal Specificity and Lawlikeness in Probabilistic Explanation // Philosophy of Science. 1968. V. 35. P. 16–33.
5. Vityaev E.E. The logic of prediction // Mathematical Logic in Asia 2005 : Proceedings of the 9th Asian Logic Conference / eds. Goncharov S.S., Downey R. and Ono. H. Novosibirsk : World Scientific, 2006. P. 263–276.
6. Смердов С.О., Витяев Е.Е. Синтез логики, вероятности и обучения: формализация предсказания // Сибирские электронные математические известия. Т. 6. Новосибирск : Ин-т математики им. С.Л. Соболева СО РАН, 2009. С. 340–365.
7. Витяев Е.Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов. Новосибирск : Новосибир. гос. ун-т.; 2006. 293 с.
8. Витяев Е.Е., Перловский Л.И., Ковалерчук Б.Я., Сперанский С.О. Вероятностная динамическая логика мышления // Нейроинформатика. 2011. Т. 5, № 1. С. 1–20.
9. Судakov К.В. Общая теория функциональных систем М. : Медицина, 1984. С. 222.
10. Витяев Е.Е. Принципы работы мозга, содержащиеся в теории функциональных систем П.К. Анохина и теории эмоций П.В. Симонова // Нейроинформатика. 2008. Т. 3, № 1. С. 25–78.
11. Бернштейн Н.А. Биомеханика и физиология движений. // Избранные психологические труды / Н.А. Бернштейн. Москва; Воронеж, 1997. С. 605.
12. Демин А.В., Витяев Е.Е. Логическая модель адаптивной системы управления // Нейроинформатика. 2008. Т. 3, № 1. С. 79–107.
13. Мухомтов В.В., Хлебников С.В., Витяев Е.Е. Улучшенный алгоритм семантического вероятностного вывода в задаче 2-мерного анимата // Нейроинформатика. 2012. Т. 6, № 1. С. 50–62.
14. Анохин К.В., Бурцев М.С., Зарайская И.Ю., Лукашев А.О., Редько В.Г. Проект «Мозг анимата»: разработка модели адаптивного поведения на основе теории функциональных систем // Восьмая национальная конференция по искусственному интеллекту с международным участием : труды конференции. М. : Физматлит, 2002. Т. 2. С. 781–789.