

Российская академия наук
Институт прикладной физики РАН
Межрегиональная ассоциация когнитивных исследований
Российская ассоциация нейроинформатики
Нижегородский государственный университет им. Н.И. Лобачевского
Национальный исследовательский университет ВШЭ
Сургутский государственный университет ХМАО-Югры
Нижегородская государственная медицинская академия
ЗАО «Нижегородское агентство наукоемких технологий»

НЕЛИНЕЙНАЯ ДИНАМИКА В КОГНИТИВНЫХ ИССЛЕДОВАНИЯХ – 2015

**ТРУДЫ
IV ВСЕРОССИЙСКОЙ КОНФЕРЕНЦИИ**

Нижний Новгород
ИПФ РАН
2015

Обнаружение «естественных» понятий

Е.Е. Витяев

Институт математики им. С.Л. Соболева СО РАН, Новосибирск

Что такое «естественная» классификация

Перед тем как перейти к «естественным» понятиям, надо понять, что такое «естественная» классификация. Фактически они говорят об одном и том же, только «естественная» классификация говорит о строении внешнего мира, на котором она основана, а «естественные» понятия, исследуемые в когнитивных науках, о восприятии «естественных» классов.

Первый достаточно подробный философский анализ «естественной» классификации принадлежит Дж. Ст. Миллю [15]. Приведем его точку зрения достаточно подробно, чтобы стал понятен смысл «естественной» классификации. Кроме того, проведенный Дж. Ст. Миллем анализ достаточно полон и на нем основано последующее изложение.

Сначала нужно отделить «искусственные» классификации от «естественных».

«Возьмем любой признак, и если одни вещи обладают им, а другие не обладают, то на нем можно основать деление всех вещей на два класса; и мы действительно так поступаем, когда создаем имя, означающее этот признак. Таким образом, число возможных классов беспрельдно, и в действительности классов существует столько же, сколько есть общих имен. <...> Но если мы обратимся к... классу «животное» или «растение»... «белый» или «красный» и посмотрим, какими особенностями индивидуумы, обнимаемые данным классом, отличаются от индивидуумов, не входящих в него, то мы найдем, что в этом отношении одни классы сильно отличаются от других. Одни классы («белый» или «красный» – *прим. Е.Е.*) обнимают вещи, отличающиеся такими особенностями, которые можно перечислить, тогда как другие («животное» или «растение» – *прим. Е.Е.*) обладают столь большим количеством признаков, что их нельзя... перечислить... В некоторых классах мало или вовсе нет характеристических черт, кроме тех, которые означаются самим названием класса: белые вещи, например, не имеют никаких общих свойств, кроме белизны... Напротив, всех общих свойств животных или растений не могла исчерпать даже и целая сотня поколений... Отсюда видно, что свойства, которые мы полагаем в основу наших классов, иногда исчерпывают все общие признаки класса или в той или другой форме подразумевают их; в других случаях, напротив, мы отбираем немногие свойства не просто из большого, но из неисчерпаемого для нас числа их. <...> И вот именно те классы, которые отличаются неопределенным множеством свойств, а не только немногими определенными признаками, признавали родами и видами логики аристоте-

левской школы... в таком случае они признавали различие в разряде и называли его существенным отличием. <...> Общих свойств у всякого настоящего естественного разряда, а следовательно, и таких общих предложений, которые можно составить относительно этого разряда или которые, наверное, будут составлять впоследствии, по мере расширения нашего знания, – неопределенное и неисчерпаемое количество. А так как, согласно первому принципу естественной классификации, классы надо составлять так, чтобы образующие их предметы обладали наибольшим количеством общих свойств, то всякая такая классификация должна признавать и вбирать в себя все разрядовые различия, существующие между теми предметами, которые она должна классифицировать» [15].

Среди общих свойств естественного разряда есть как наблюдаемые, так и не наблюдаемые – скрытые признаки, которые являются, возможно, более важными. Чтобы учитывать скрытые признаки, надо находить их причинные проявления в наблюдаемых признаках. Дж. Ст. Милль по этому поводу пишет: «Таким образом, наши естественные группы часто должны основываться на неочевидных свойствах вещей, раз эти свойства важнее очевидных. Но в таких случаях необходимо должно быть некоторое другое свойство или ряд свойств, которые были бы более заметны для наблюдателя и сосуществовали бы в то же время со свойствами, на которых основана данная классификация, являясь их признаками. Так, например, естественная классификация животных должна основываться преимущественно на их внутреннем строении; однако было бы странно, как замечает О. Конт, если бы мы были в состоянии определить род и вид того или другого животного, только предварительно убив его. На этом основании предпочтения среди зоологических классификаций заслуживает, вероятно, классификация де Блэнвиля, основанная на различиях во внешних покровах животных» [15].

Таким образом, «естественную» классификацию Дж. Ст. Милль определяет следующим образом: «Всего более соответствует целям научной (естественной) классификации, когда предметы соединяются в такие группы, относительно которых можно высказать наибольшее число общих предложений... предметы следовало бы классифицировать по возможности на основании таких свойств, которые служат причинами многих других или, по крайней мере, составляют их верные признаки. Из этих последних надо выбирать такие, которые были бы самыми надежными и наиболее непосредственными признаками... Но, к

сожалению, свойства, служащие причинами главных отличительных признаков классов, редко бывают в то же время способны служить для узнавания этого класса. И вместо причин нам по большей части приходится выбирать те или другие из наиболее бросающихся в глаза их следствий – такие следствия, которые могут служить признаками как других следствий, так и самой причины» [15].

На основании понятия «естественного» класса Дж. Ст. Милль определяет также понятие «образа» класса, которое является предтечей «естественных» понятий, возникающих в когнитивной психологии и рассмотренных далее.

«Естественные группы не только не менее всяких искусственных классов определяются признаками, но образуются именно ввиду и на основании признаков. Однако при этом принимаются во внимание не только признаки, безусловно общие для всех включаемых в группу предметов, но вся совокупность тех признаков, из которых все встречаются в большинстве этих предметов, а большинство – во всех. Вследствие этого наше понятие о классе – тот образ, которым этот класс представлен в нашем уме, – есть понятие о некотором образце, обладающем всеми признаками данного класса, естественнее же всего о таком образце, который, обладая всеми этими признаками в самой высокой степени, в какой они когда-либо встречаются, вследствие этого всего яснее и резче может указать нам, в чем состоят эти признаки. К этому-то мерилу – не вместо определения, а в пояснение его – мы обыкновенно и с удобством обращаемся с целью определить, принадлежит ли тот или другой индивидуум или вид к данному классу или нет» [15].

Рассуждения Дж. Ст. Милля были подтверждены естествоиспытателями, строившими «естественные» классификации. Например, В. Уэвель пишет, что в «естественных» классах мы можем сделать наибольшее число общих предложений: «Чем больше общих утверждений об объектах дает возможность сделать классификация, тем она естественней» [7]. О неисчерпаемом количестве общих свойств у «естественных» классов пишет Л. Рутковский: «Чем в большем числе существенных признаков сходны сравниваемые предметы, тем вероятнее их одинаковость и в других отношениях» [9]. Аналогичное высказывание делает Е.С. Смирнов: «Таксономическая проблема заключается в “индикации”: от бесконечно большого числа признаков нам нужно перейти к ограниченному их количеству, которое заменило бы все остальные признаки» [10].

В результате была сформулирована проблема определения «естественных» классификаций, которая до сих пор обсуждается в литературе [5, 6, 8, 11, 12, 22]. Однако, с нашей точки зрения, до сих пор нет достаточно адекватного формального определения «естественной» классификации.

Принципы категоризации в когнитивных науках

В работах Э. Рош [16–20] на основании проведенных экспериментов сформулированы принципы категоризации «естественных» категорий, подтверждающие высказывания Дж. Ст. Милля и естествоиспытателей.

Первый принцип – «когнитивной экономии» – состоит в том, что организм хочет получить от категорий как можно больше информации об окружении, используя как можно меньше ресурсов. С одной стороны, это позволяет предсказывать как можно больше свойств по отдельным свойствам, а с другой стороны, формировать большое число категорий с наилучшим разделением между собой.

Второй принцип – «структурности воспринимаемого мира» – дополняет первый и состоит в том, что воспринимаемый мир не является неструктурированным множеством равновероятно встречающихся свойств, наоборот, объекты воспринимаемого мира имеют высоко коррелированную структуру. При этом надо иметь в виду, что речь идет о воспринимаемом мире, а не о метафизическом мире «в себе». Собака и человек видят мир по-разному.

Понятно, что первый принцип не возможен без второго – когнитивная экономия не возможна без структурированности мира.

Непосредственно воспринимаемые объекты (basic objects) – информационно богатые связки наблюдаемых и функциональных свойств, которые образуют естественную разрывность, создающую категоризацию. Эти связки формируют «прототипы» объектов классов (образец у Дж. Ст. Милля).

«Categories can be viewed in terms of their clear cases if the perceiver places emphasis on the correlational structure of perceived attributes... By prototypes of categories we have generally meant the clearest cases of category membership» [19, 20]. «Rosch and Mervis (1975) have shown that the more prototypical of a category a member is rated, the more attributes it has in common with other members of the category and the fewer attributes in common with members of the contrasting categories» [17–18].

В дальнейшем теория «естественных» понятий Э. Рош получила название прототипической теории понятий (prototype theory). Основные ее черты описываются следующим образом: «прототипическая точка зрения придерживается предположения, что есть некоторое общее множество свойств членов категории, не предполагающее, что все члены категории обладают всеми этими свойствами. Наоборот, предполагается вероятностный процесс сходства: члены категории имеют больше общих свойств, возможно взвешенных по важности, с прототипом данной категории, чем с прототипами других категорий» [21].

К сожалению, в такой формулировке исчезают принципы категоризации, сформулированные Э. Рош, говорящие об отражении в «естественных» понятиях высококоррелированной структуры внешнего мира.

Выдвинутые затем математические модели формирования понятий, основанные на таком понимании прототипической теории, также не отражают высококоррелированную структуру мира.

Формализация «естественных» классификаций и «естественных» категорий

Предлагаемая нами формализация «естественных» классификаций и «естественных» категорий прямо основана на высокой коррелированности структуры внешнего мира. Более того, основным предположением и механизмом построения «естественной» классификации является обнаружение всех возможных причинных связей, вытекающих из этой коррелированности [1]. Тогда, если обнаружение причинных связей формализовать специальным семантическим вероятностным выводом и представить набором правил, а высококоррелированную структуру – замыканием всех предсказаний, осуществляемых этими причинными связями, на себя, то получатся неподвижные точки предсказаний, формирующие прототипы «естественных» классов.

Математический аппарат неподвижных точек предсказаний по причинным связям основан на следующих новых и нетривиальных математических результатах:

1) семантический вероятностный вывод, обнаруживающий причинные связи, решает также проблему статистической двусмысленности – нами доказано, что он обнаруживает максимально специфические причинные правила, предсказывающие без противоречий [2, 24];

2) семантический вероятностный вывод может быть представлен как формальная модель нейрона, а неподвижные точки, как клеточные ансамбли [25];

3) неподвижные точки являются вероятностным обобщением формальных понятий [3, 13, 26], исследуемых в анализе формальных понятий [14];

4) можно доказать, что неподвижные точки предсказаний для правил с отрицанием логически непротиворечивы [4, 27].

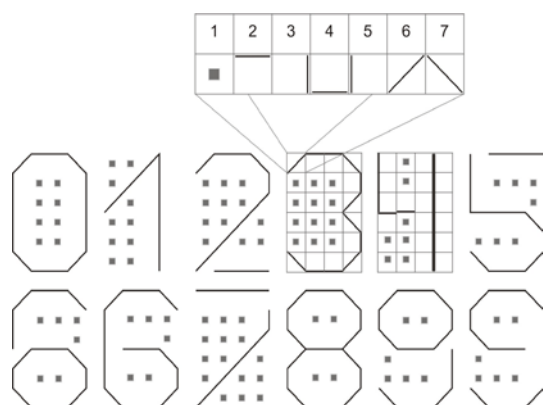


Рис. 1. Кодировка цифр

Проиллюстрируем неподвижные точки компьютерным экспериментом по формированию «естественных» классов и понятий закодированных цифр.

Закодируем цифры, как показано на рис. 1. Сформируем обучающее множество, состоящее из 360 перетасованных цифр (12 цифр рис. 1 продублированных в 30 экземплярах без указания, где какая цифра). На этом множестве семантическим вероятностным выводом обнаружено 55089 закономерностей – общих утверждений об объектах, о которых говорит Дж. Ст. Милль и В. Уэвель.

По этим закономерностям обнаружено ровно 12 неподвижных точек, которые в точности соответствуют нашим цифрам, хотя алгоритму ничего не известно о содержании объектов, кроме значений признаков.

Пример неподвижной точки для цифры 6 приведен на рис. 2. Рассмотрим, что представляет собой неподвижная точка. Пронумеруем признаки цифр, как указано в таблице.

Кодировка полей цифр

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16
17	18	19	20
21	22	23	24

Первая закономерность цифры 6 рис. 2, представленная в первом прямоугольнике после фигурной скобки, говорит, что, если в квадрате 13 (табл.) стоит признак 6 (обозначим это как 13-6), то в квадрате 3 должен стоять признак 2 (обозначим как 3-2). Предсказываемый признак обозначается точечной линией. Запишем эту закономерность как $13-6 \Rightarrow 3-2$. Нетрудно проверить по цифрам, что эта закономерность действительно выполнена на всех цифрах. Вторая закономерность говорит, что из признака 9-5 и отрицания значения 5 первого признака $\neg 1-5$ (первый признак не должен быть равен 5) следует признак 4-7.

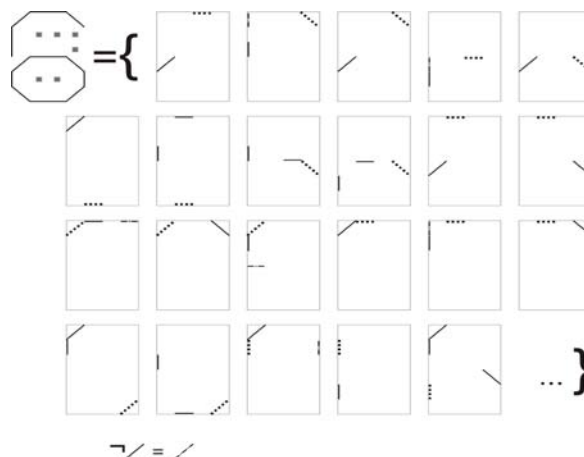


Рис. 2. Неподвижная точка цифры 6

Отрицание значения некоторого признака обозначается пунктирной линией, как показано в нижней части рис. 2. В результате получаем закономерность $9-5 \&-(1-5) \Rightarrow 4-7$. Последующие 3 закономерности в первой строке цифры 6 на рис. 2 будут соответственно закономерности $(13-6 \Rightarrow 4-7)$, $(17-5 \&-(13-5) \Rightarrow 4-7)$, $(13-6 \Rightarrow 16-7)$.

На рис. 2 видно, что закономерности и признаки цифры 6 образуют неподвижную точку – взаимно предсказывают друг друга, замыкая взаимные предсказания признаков на себя. Заметим, что при этом закономерности, используемые в неподвижной точке, выполнены на всех цифрах, а сама неподвижная точка выделяет только одну цифру. Поэтому цифры выделяются не закономерностями самими по себе, а их взаимосвязью.

Неподвижная точка автоматически обнаруживает паттерн – набор признаков, составляющий неподвижную точку, который является прототипом у Э. Рош и образом у Дж. Ст. Милля. Программа не знает заранее, какие сочетания признаков максимально коррелируют между собой (взаимно предсказываются причинными связями) и составляют паттерн.

Работа поддержана грантом РФФИ № 15-07-03410-а.

Литература

1. Витяев Е.Е. Классификация как выделение групп объектов, удовлетворяющих разным множествам согласованных закономерностей // Анализ разнотипных данных. Новосибирск, 1983. С. 44–50. (Выч. сист. ; вып. 99).
2. Витяев Е.Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов. Новосибирск : НГУ, 2006. С. 293.
3. Витяев Е.Е., Демин А.В., Пономарёв Д.К. Вероятностное обобщение формальных понятий // Программирование. 2012. Т. 38, № 5. С. 219–230.
4. Витяев Е.Е., Мартынович В.В. Вероятностные формальные понятия на контекстах с отрицаниями // Информационные технологии в гуманитарных исследованиях. Новосибирск, 2014. Вып. 19. С. 5–20.
5. Забродин В.Ю. О критериях естественной классификации // НТИ. Сер. 2. 1981. № 8.
6. Кожара В.Л. Функции классификации // Теория классификаций и анализ данных. Новосибирск, 1982.
7. Мейен С.В., Шрейдер С.А. Методологические аспекты теории классификаций // Вопросы философии. 1976. № 12.
8. Розова С.С. Классификационная проблема в современной науке. Новосибирск : Наука, 1986. 224 с.
9. Рутковский Л. Элементарный учебник логики. СПб, 1884.
10. Смирнов Е.С. Конструкция вида таксономической точки зрения // Зоол. журн. 1938. Т. 17, № 3. С. 387–418.
11. Субботин А.Л. Классификация. М. : ИФ РАН, 2001.
12. Шрейдер С.А. Систематика, типологии, классификация // Теория и методология биологических классификаций. М. : Наука, 1983.
13. Demin A., Ponomarev D., Vityaev E. Probabilistic Concepts in Formal Contexts // Preliminary Proceedings of the Ershov Informatics Conference PSI Series, 8-th Edition (June 27 – July 1, 2011, Novosibirsk), Novosibirsk, 2011. P. 29–38.
14. Ganter B. Formal Concept Analysis: Methods, and Applications in Computer Science. TU Dresden, Germany, 2003.
15. Mill J.S. System of Logic, Ratiocinative and Inductive. L., 1983.
16. Rosch E.H. Natural categories // Cognitive Psychology. 1973. V. 4, Is. 3. P. 328–350.
17. Rosch E., Mervis C.B. Family resemblances. Studies in the internal structure of categories // Cognitive Psychology. 1975. V. 7, Is. 4. P. 573–605.
18. Rosch E., Mervis C. B., Gray W.D., Johnson D.M., Boyes-Braem P. Basic objects in natural categories // Cognitive Psychology. 1976. V. 8. P. 382–439.
19. Rosch E., Lloyd B.B. (eds) Cognition and categorization. Hillsdale, N.J. : Lawrence Erlbaum, 1978. P. 27–48.
20. Rosch E. Principles of Categorization // Rosch E., Lloyd, B.B. (eds) Cognition and Categorization. Lawrence Erlbaum Associates, 1978. P. 27–48.
21. Ross B.H., Taylor E.G., Middleton E.L., Nokes T.J. Concept and Category Learning in Humans // Learning and Memory: A Comprehensive Reference. V. 2 : Cognitive Psychology of Memory. Oxford : Elsevier, 2008. P. 535–556.
22. Scientific Discovery / website. URL: <http://math.nsc.ru/AP/ScientificDiscovery/>
23. The Nature of Classification : Relationships and Kinds in the Natural Sciences. Palgrave Macmillan, 2013. 208 p.
24. Vityaev E. The logic of prediction // Mathematical Logic in Asia : Proceedings of the 9th Asian Logic Conference / edited by S.S. Goncharov, R. Downey, H. Ono. Singapore : World Scientific, 2006. P. 263–276.
25. Vityaev E.E. A formal model of neuron that provides consistent predictions // Biologically Inspired Cognitive Architectures 2012 : Proceedings of the Third Annual Meeting of the BICA Society. Heidelberg : New York ; Dordrecht ; London : Springer, 2013. P. 339–344. (Advances in Intelligent Systems and Computing. V. 196).
26. Vityaev E.E., Demin A.V., Ponomaryov D.K. Probabilistic Generalization of Formal Concepts // Programming and Computer Software. 2012. V. 38, № 5. P. 219–230.
27. Vityaev E.E., Martinovich V.V. Probabilistic Formal Concepts with Negation // Perspectives of System Informatics. 2015. P. 385–399. (Lecture Notes in Computer Science ; vol. 8974). DOI: 10.1007/978-3-662-46823-4_31.