

УДК 681.142/155

ОБ ОПРЕДЕЛЕНИИ ПОНЯТИЯ "ЗАКОНОМЕРНОСТЬ"

Н.Г.Загоруйко

Как на Первом, так и на Втором Всесоюзных симпозиумах по проблеме "Машинные методы обнаружения закономерностей" (МОЗ-1, 1976, Новосибирск; МОЗ-2, 1969, Рига) определению понятия "закономерность" было уделено много внимания [1]. В общих словах итоги дискуссий по этому вопросу сводятся к следующему: закономерность (R) есть некоторое гипотетическое утверждение (h) об определенных свойствах изучаемых объектов и само обладающее определенными характеристиками (L): $R = h/L$.

Эмпирическая гипотеза h определяется как набор, состоящий из следующих четырех элементов:

$W = \{a_1, a_2, \dots, a_n\}$ - непустое (конечное или бесконечное) множество объектов, относительно которых высказывается гипотеза;

$v = \{P_1^1, P_2^1, \dots, P_n^1\}$ - словарь, т.е. непустое конечное множество символов (m_1 -местных предикатов), с помощью которых фиксируются в протоколах pr результаты наблюдений над объектами;

$O = \{I_1, I_2, \dots, I_1, \dots, I_n\}$ - средства наблюдения, с помощью которых устанавливаются значения истинности предикатов из v ;

$T = \{\tau_1, \tau_2, \dots, \tau_1, \dots, \tau_n\}$ - "тестовый алгоритм" - непустое конечное множество процедур, с помощью которых принимается решение о том, согласуется ли некоторый результат эксперимента (т.е. протекол pr) с тем, что утверждает гипотеза, или не согласуется.

Гипотезы $h = \langle W, v, O, T \rangle$ можно сравнивать не только по тому, о чем они говорят, но и по тому, что и как они говорят об этом. Внутри одной содержательной области гипотезы можно сравнивать по следующим характеристикам из набора L :

1. Потенциальная опровержимость Q . Чем большую долю $\bar{N}$ из всех возможных неизоморфных протеколов

И гипотеза предполагает невозможными, тем больше ее потенциальный риск быть опровергнутой в ходе проверки.

Неопровержимая гипотеза - это та, которая ничего не запрещает, у нее $\bar{N} = 0$ и $Q = \frac{\bar{N}}{N} = 0$. Наиболее сильные гипотезы запрещают все, кроме того, что подчиняется конкретному "закону природы", их характеристика Q имеет максимально возможное значение. Основное назначение гипотез - делать предсказания, а предсказывает гипотеза тем точнее, чем больше возможностей она запрещает. Этим объясняется, почему гипотеза с более высоким значением Q считается более полезной, более "сильной".

Если каждый из n предикатов может входить в протокол pg в одном из 1 значений истинности, то $N = 1^n$.

Следует отметить, что сравнивать по характеристике Q можно лишь те гипотезы, которые имеют одинаковое число возможных неизоморфных протоколов N .

2. Подтвержденность P . Так как гипотезы обычно содержат утверждения о бесконечных множествах объектов W ("для всех пар заряженных тел"... и т.д.), то доля M экспериментов, которые уже проводились и не опровергли гипотезу h , по отношению к бесконечному числу M возможных экспериментов, отличающихся хотя бы одним объектом a_i из набора W , будет всегда бесконечно малой. Однако человек постоянно пользуется гипотезами, полная проверка которых невозможна, и по конечному числу M упорядочивает гипотезы по степени подтвержденности, например, в такой шкале: "непроверенная", "маловерифицированная", "хорошо проверенная", "многочисленно и безупречно проверенная".

По-видимому, изучение эксперта, дающего такие оценки, как измерительного прибора, позволит формализовать вычисление характеристики P для гипотез с бесконечным базисом W .

3. Объясненность E . Если мы знаем не только, что возможно и что невозможно, но еще и почему это возможно или невозможно, то в этом случае наше доверие к гипотезе возрастает. Р.Фейнман [2] считает объясненность одной из наиболее важных характеристик эмпирических гипотез: только тогда мы можем быть уверены в справедливости физического закона, если мы можем его объяснить с помощью нескольких моделей.

Степень объясненности E является функцией числа последовательных ответов на вопросы типа "что?", "как?" и "почему?": "Что это?", "Почему именно это, а не другое?", "Как это устроено (работает, происходит)?", "Почему именно так, а не иначе?" и т.д.

Замечено [3], что человека обычно устраивает последовательность — ность ответов на два-три таких вопроса. В [3] дается формальный аппарат для описания процесса углубления понимания через последовательность объяснений.

4. М о щ н о с т ь D . Мощность объектного базиса гипотезы измеряется числом возможных различных экспериментальных ситуаций, если различными мы будем считать ситуации, отличающиеся хотя бы одним объектом из участвующих в эксперименте. Если гипотетическое утверждение касается свойств любых q из m объектов множества W , то очевидно, что $D = C_m^q$.

В частности, если гипотеза говорит нечто о любом отдельно взятом объекте (например: "на любое тело, погруженное в воду, действует выталкивающая сила"... и т.д.), то $D = m$. Из-за того, что m обычно бесконечно, сравнивать между собой по величине D можно лишь такие гипотезы h_1 и h_2 , области задания (объектные базисы) которых находятся в отношении включения. Если $W_1 \subset W_2$, то естественно считать гипотезу h_2 более сильной по характеристике D , чем h_1 .

5. П р о с т о т а S . Одно и то же эмпирическое содержание может быть выражено гипотезами, отличающимися сложностью своих формулировок. Согласно "гипотезе простоты" [4], предпочтение отдается более простым формулировкам. Одному из возможных способов измерения сложности гипотез посвящена работа [5] в данном сборнике.

6. Можно указать на еще одну характеристику, связанную с особенностями формулировок гипотез, — к р а с о т у или из я щ е с т в о (В) г и п о т е з.

В истории науки известны примеры, подтверждающие важную эвристическую роль "гипотезы красоты" [6,7]. Сейчас было бы трудно рассчитывать на какой-нибудь объективный метод измерения изящества теоретической конструкции, кроме метода экспертных оценок.

Исходные гипотезы, с которых обычно начинаются исследования, имеют минимальные (в своей истории) значения всех этих характеристик:

$$L_{min} = \{Q_{min}, P_{min}, E_{min}, D_{min}, S_{min}, B_{min}\}.$$

Эти гипотезы—"претенденты" мало или совсем пока не подтверждались, плохо объясняются, формулируются сложно и неряшливо. Гипотезы-претенденты служат сырьем, стимулирующим целенаправленные исследования в некоторой предметной области.

Конечный же продукт науки - гипотезы-"законы природы"- обладают наибольшими значениями своих характеристик ($L_{\max}$). Они рискованны, но бесспорно подтверждены, глубоко объяснены, просты и изящны и т.д.

В промежутке между этими крайностями лежит область утверждений h , которые достаточно рискованны, но уже имеют некоторое число подтверждений и некоторое объяснение, их формулировки еще не совершенны, но уже не такие "сырые", как раньше.

Гипотезы с такими "средними" характеристиками L^* и относятся к "закономерностям".

В отличие от гипотез-"законов" их продолжают проверять и усиливать, но, в отличие от гипотез-"претендентов", уже начинают использовать для предсказаний.

Изучение процессов усиления гипотез этого класса, перехода их из гипотез-претендентов в гипотезы-закономерности и составляет основное ядро проблемы МОЗ - машинных методов обнаружения закономерностей.

Из сказанного выше ясно, что задачу определения понятия "закономерность" нельзя считать полностью решенной.

На пути к ее решению предстоит разработать объективные методы измерения значений всех L -характеристик эмпирических гипотез.

Л и т е р а т у р а

1. Машинные методы обнаружения закономерностей. Материалы Всесоюзного симпозиума.-Новосибирск, Наука, 1976, -160 с.
2. ФЕЙНМАН Р. Природа физических законов.-М.: Мир, 1967.-180 с.
3. ЗАГОРУЙКО Н.Г., СВИРИДЕНКО Д.И. Формализация процесса углубления понимания.-В кн.: Эмпирическое предсказание и распознавание образов. (Вычислительные системы, вып. 67.) Новосибирск, 1976, с. 87-92.
4. МАМЧУР Е.А. Проблема выбора теории.-М.: Наука, 1975.-150 с.
5. САМОХВАЛОВ К.Ф. Простота эмпирических теорий. -Настоящий сборник, с.12-14.
6. УОТСОН Д.Д. Двойная спираль.-М.: Мир, 1966.-130 с.
7. МИГДАЛ А. О психологии научного творчества.-Наука и жизнь, 1976, №3, с.5-15.

Поступила в ред.-изд.отд.
12 ноября 1979 года