ЭМПИРИЧЕСКОЕ ПРЕДСКАЗАНИЕ И РАСПОЗНАВАНИЕ ОБРАЗОВ
(Вычислительные системы)
1979 год Выпуск 79

УДК 518.74

ОБНАРУЖЕНИЕ ЗАКОНОМЕРНОСТЕЙ, ВЫРАЖЕННЫХ УНИВЕРСАЛЬНЫМИ ФОРМУЛАМИ

Е.Е. Витяев

В работе [I] доказано, что при выполнении определенных требований на эксперимент существует совокупность универсальных формул W, одновременно выполнимых на тех и только тех конечных моделях, которые являются моделями серии экспериментов. В [2] приведен метод обнаружения универсальных закономерностей. В настоящей работе показывается, в каком смысле метод [2] обнаруживает на подходящем обучающем материале любую закономерность, выраженную универсальной формулой, и, значит, в силу [I] - любую конечно выразимую экспериментальную зависимость.

ПРЕДПОЛОЖЕНИЕ. Пусть N - некоторая не пустая совокупность объектов, которую назовем генеральной совокупностью. Предположим, что серии экспериментов проводятся в заданной экспериментальной ситуации, удовлетворяющей требованиям, сформулированным в [I], над случайно выбранными из N с помощью некоторой фиксированной процедуры, объектами. Каждый случайно выбранный набор объектов из N имеет не нулевую вероятность.

В силу этого предположения каждая модель серии экспериментов $S = \langle M;(Q_i)\rangle$ имеет некоторую не нулевую вероятность. Множество универсальных формул W одновременно выполнимо только на таких моделях.

ОПРЕДЕЛЕНИЕ I. Импликативным следствием $W \overset{M}{\vdash} \Phi$ совокупности формул W назовем формулу $\Phi = \forall x_1,\dots,x_n \ (Q_1^{z_1} \& \dots \& Q_n^{z_n} \Rightarrow Q_0^{z_0})$, которая является следствием W и перестает быть им после удаления одного из символов отношений $Q_1^{z_1},\dots,Q_n^{z_n}$ или после замены символа отношения $Q_0^{z_0}$ на символ Л (ложь). Величины $z_i = 0,I$; $i = 0,I,\dots,n$, задают наличие отрицаний.

ЛЕММА. Множество универсальных формул W логически эквивалентно некоторой совокупности своих импликативных следствий.

ДОКАЗАТЕЛЬСТВО. В [2] доказано, что W эквивалентно некоторой совокупности формул U вида $\Phi = \forall x_1, \dots, x_n (Q_1^{z_1} \& \dots \& Q_n^{z_n} \Rightarrow Q_0^{z_0})$. Преобразуем множество U следующим образом. Если формула $\Phi \in U$ остается следствием U после удаления в ней одного из символов отношений в посылке или после замены символа отношения $Q_0^{z_0}$ на символ Л, то заменим ее в U на формулу Φ', полученную из Φ только что указанным способом. Так как каждая формула Φ' логически сильнее соответствующей формулы Φ и в то же время она является следствием U, то множество U перейдет в результате такой замены в эквивалентное себе множество. Проведем все возможные преобразования такого вида. В результате мы получим множество импликативных следствий, эквивалентное исходному множеству W.

ОПРЕДЕЛЕНИЕ 2. Формулу вида $\Phi = \forall x_1, \dots, x_n \ (R_1 \& \dots \& R_n \Rightarrow R_0)$ или, в других обозначениях, $\forall x_1, \dots, x_n (D^{\&} \Rightarrow R_0)$, $R_i = Q_i^{z_i}$, $i = 0, 1, \dots, n$; $D = \{R_1, \dots, R_n\}$; $D^{\&}$ – конъюнкция всех отношений из D, будем называть несводимой закономерностью, если для любого подмножества $Q \subset D$, $Q \neq D$, имеют место соотношения $P(D^{\&}) \neq 0$ и $P(R_0 / D^{\&}) > P(R_0/(D \setminus Q)^{\&})$ (здесь P – вероятности, индуцируемые случайным выбором наборов $\langle a_1, \dots, a_n \rangle$ объектов из генеральной совокупности N).

ТЕОРЕМА. Выводимая из W формула $\Phi = \forall x_1, \dots, x_n \ (R_1 \& \dots \& R_n \Rightarrow R_0)$ является импликативным следствием $W \overset{И}{\vdash} \Phi$ тогда и только тогда, когда она является несводимой закономерностью.

ДОКАЗАТЕЛЬСТВО. Из $W \vdash \Phi$ следует, что на моделях серий экспериментов $P(R_0 / D^{\&}) = 1$. Если Φ является несводимой закономерностью, то это эквивалентно тому, что для любого $Q \subset D$, $Q \neq D$, $P(R_0/(D \setminus Q)^{\&}) < 1$ и $P(D^{\&}) > 0$. Эти неравенства, в свою очередь, эквивалентны тому, что для любого Q формулы $\forall x_1, \dots, x_n ((D \setminus Q)^{\&} \Rightarrow R_0)$ и $\forall x_1, \dots, x_n (D^{\&} \Rightarrow Л)$ ложны на некоторых моделях серий экспериментов. Это возможно тогда и только тогда, когда они не являются следствиями множества формул W . Что, в свою очередь, эквивалентно тому, что Φ – импликативное следствие W_0.

Для обнаружения несводимых закономерностей и проверки соответствующих вероятностных неравенств в методе [2] используется точный критерий независимости Фишера, который при заданном доверительном уровне на подходящем обучающем материале может обнаружить нужную закономерность.

Если множество W конечно, то существует универсальная формула Φ, эквивалентная W и выражающая экспериментальную зависимость. В силу леммы и теоремы для Φ существует эквивалентная ей конечная совокупность несводимых закономерностей, которая может быть обнаружена на подходящем обучающем материале. В этом смысле метод [2] обнаруживает любую закономерность, выраженную универсальной формулой.

Если нам не известна мощность множества W и дана формула Φ, истинная на всех моделях серий экспериментов, то $W \vdash \Phi$ и, в силу леммы и теоремы, также существует конечная совокупность несводимых закономерностей V, из которой следует Φ, $V \vdash \Phi$.

Литература

1. ВИТЯЕВ Е.Е. Закономерности в языке эмпирических систем.-В кн.: Эмпирическое предсказание и распознавание образов. (Вычислительные системы, вып. 76.) Новосибирск, 1978, с. 3-14.

2. ВИТЯЕВ Е.Е. Метод обнаружения закономерностей и метод предсказания. -В кн.: Эмпирическое предсказание и распознавание образов. (Вычислительные системы, вып. 67.) Новосибирск,1979,с.54-68.

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