МАШИННЫЙ АНАЛИЗ СЛОЖНЫХ СТРУКТУР
(Вычислительные системы)
1986 год Выпуск 118

УДК 51.16

ОБ ОДНОМ СВОЙСТВЕ МЕТОДОВ ИНДУКЦИИ НАД СТАНДАРТНЫМИ ЭМПИРИЧЕСКИМИ ТЕОРИЯМИ

А.С.Нудельман

Основным предметом методологического анализа проблемы индукции, начатого в [1,2] и продолженного в [3,4], является фиксация методологически обоснованных свойств методов индукции и поиск методов, которые обладали бы такими свойствами. В данной работе формулируется и обосновывается новое свойство методов индукции, названное монотонностью, и, кроме того, выясняется наличие (отсутствие) такого свойства у некоторых известных методов индукции.

1. Будем использовать принятые в [2,4] следующие обозначения и соглашения:

E - совокупность всех конечных (непустых) множеств эмпирических объектов во всех возможных (эмпирических) мирах;

α - фиксированный счетный алфавит символов;

I^A - именующее отображение (взаимно-однозначное отображение множества $A \in E$ в α). Если $A' \subseteq A$, то $I^A \restriction A'$ обозначает ограничение отображения I^A на множестве A';

v - словарь (конечная предикатная сигнатура);

M^v - класс всех конечных моделей сигнатуры v, носители которых - множества из E. Если $m \in M^v$, то $|m|$ обозначает носитель модели m;

$pr^v(m, I^{|m|})$ - протокол в словаре v (диаграмма модели $m \in M^v$ при именующем отображении $I^{|m|}$);

pr^v - протокол в словаре v (диаграмма некоторой модели $m \in M^v$ при некотором именующем отображении $I^{|m|}$). Выражение $B(pr^v)$ обозначает базис протокола pr^v, т.е. множество всех индивидных констант (символов из α), участвующих в записи этого протокола. Мощность множества $B(pr^v)$ называется мощностью протокола pr^v и обозначается через $\bar{\bar{B}}(pr^v)$. Если множество $B' \subseteq B(pr^v)$ непусто, то выражение $pr^v \restriction B'$ обозначает ограничение протокола pr^v на множестве B' , т.е. протокол, который может быть получен из pr^v удалением всех элементов, содержащих символы из дополнения $B(pr^v) \setminus B'$. О протоколах pr_1^v и pr_2^v говорится, что они изоморфны (пишется: $pr_1^v \simeq pr_2^v$), если и только если они могут быть сделаны равными взаимно-однозначной переименовкой базиса одного из них;

obs^v - инструкция (в словаре v) о том, чем и как проводить наблюдения эмпирических объектов. Если инструкция obs^v применима для наблюдения множества $A \in E$ (символически: $obs^v(A) \neq \emptyset$), то результатом наблюдения этого A в соответствии с инструкцией obs^v будет модель $obs^v(A) \in M^v$,а записью результата такого наблюдения при выбранном I^A - протокол $pr^v(obs^v(A), I^A)$.Инструкция obs^v называется стандартной тогда и только тогда, когда для всякого $A \in E$ такого, что $obs^v(A) \neq \emptyset$, всякого отображения I^A и всякого $A' \in E$ такого, что $A' \subseteq A$, выполняется следующее: если наблюдение множества A в соответствии с инструкцией obs^v состоялось и записью этого наблюдения является протокол $pr^v(obs^v(A), I^A)$,то состоялось наблюдение в соответствии с obs^v множества A' и записью такого наблюдения является протокол $pr^v(obs^v(A'), I^A \restriction A')$, который совпадает с $pr^v(obs^v(A), I^A) \restriction ran(I^A \restriction A')$;

T^v - тестовый алгоритм в словаре v (отображение множества всех протоколов в словаре v в $\{0, 1\}$, причем если протоколы pr_1^v и pr_2^v изоморфны, то $T^v(pr_1^v) = T^v(pr_2^v)$). Класс всех тестовых алгоритмов обозначается через τ . Тестовый алгоритм T^v называется стандартным, если и только если для любого протокола pr^v такого, что $T^v(pr^v) = 1$, и любого непустого подмножества B' базиса $B(pr^v)$ имеет место $T^v(pr^v \restriction B') = 1$. Пара $\langle T^v, pr^v \rangle$ называется допустимой тогда и только тогда, когда $T^v(pr^v) = 1$. Класс всех допустимых пар обозначается через π ;

h - эмпирическая теория (тройка $\langle v, obs^v, T^v \rangle$, где v - словарь теории h, obs^v - инструкция теории h и T^v - тестовый алгоритм теории h). Эмпирический смысл теории $h = \langle v, obs^v, T^v \rangle$ определяется следующим соглашением: если инструкция obs^v применима для наблюдения множества $A \in E$, то считается, что результат наблюдения множества A , проведенного в соответствии с инструкцией obs^v, согласуется с теорией h , если $T^v(pr^v(obs^v(A))) = 1$;

и считается, что результат такого наблюдения опровергает теорию h, если $T^V(pr^V(obs^V(A))) = 0$;

f – метод индукции, т.е. функция, ставящая в соответствие каждой согласованной тройке $\langle h_0, pr_0, A_0 \rangle$ некоторую эмпирическую теорию h_1. Заметим, что тройка $\langle h_0, pr_0, A_0\rangle$ согласована тогда и только тогда, когда $A_0 \in E$, и если $h_0 = \langle v, obs^V, T_0^V\rangle$, то $obs^V(A_0) \neq \emptyset$, $pr_0 = pr^V(obs^V(A_0))$ и $T_0^V(pr_0) = 1$.

2. Эмпирическую теорию $h = \langle v, obs^V, T^V\rangle$ будем называть стандартной тогда и только тогда, когда obs^V – стандартная инструкция и T^V – стандартный тестовый алгоритм. Понятие стандартной эмпирической теории является (конечно, предположительно) экспликацией понятия реальной естественнонаучной теории.

Метод индукции f будем называть методом над стандартными эмпирическими теориями, если и только если: а) метод f применяется для усиления только стандартных теорий, т.е. областью применения метода f является класс всех согласованных троек, содержащих стандартные теории; б) для всякой согласованной тройки $\langle h_0, pr_0, A_0\rangle$, где теория h_0 стандартна, эмпирическая теория $f(h_0, pr_0, A_0)$ тоже стандартна.

3. Сформулируем и обоснуем новое свойство – свойство монотонности – методов индукции. При этом будем иметь в виду, что всякий метод индукции f удовлетворяет исходному свойству C_f (пп. 1–3 взяты из [2, с.32–33]): для произвольной согласованной тройки $\langle h_0, pr_0^V, A_0\rangle$ и произвольной эмпирической теории h_1, если $h_0 = \langle v, obs^V, T_0^V \rangle$ и $h_1 = f(h_0, pr_0^V, A_0)$, то

C_f-1) $h_1 = \langle v, obs^V, T_1^V \rangle$ и $T_1^V = ind_f(T_0^V, pr_0^V)$, где ind_f – некоторое отображение из π в τ;

C_f-2) $T_1^V(pr_0^V) = 1$;

C_f-3) для любого протокола pr^V, если $T_0^V(pr^V) = 0$, то $T_1^V(pr^V) = 0$;

C_f-4) для любого протокола pr^V, если $pr^V \simeq pr_0^V$, то $ind_f(T_0^V, pr^V) = T_1^V$.

Монотонность методов индукции будет характеризовать одну из особенностей реальной эпистемологической практики, а именно: в рамках принятой эпистемологической установки (метода индукции) теория не изменяется существенным образом, пока новые экспериментальные данные согласуются с этой теорией (такие данные могут обусловить только уточнение (усиление) теории). Заметим, что свойство C_f-3 характеризует эту же особенность эпистемологической практики.

Существенным здесь является то обстоятельство, что обоснованность нового свойства ограничена только методами индукции над стандартными эмпирическими теориями. Такое ограничение представляется естественным, поскольку реальная эпистемологическая практика имеет дело, по-видимому, только со стандартными теориями.

Свойство монотонности. Для всяких тестовых алгоритмов T_0^v, T_1^v, T_2^v в словаре v и всяких протоколов pr_1^v, pr_2^v в словаре v, если $T_0^v(pr_1^v) = 1$, $T_1^v = ind_f(T_0^v, pr_1^v)$, $T_0^v(pr_2^v) = 1$, $T_2^v = ind_f(T_0^v, pr_2^v)$ и если $pr_1^v \subset pr_2^v$ и $T_1^v(pr_2^v) = 1$, то для любого протокола pr^v в словаре v выполняется $T_1^v(pr^v) = 0 \to T_2^v(pr^v) = 0$.

Обоснование. Покажем методологическую неприемлемость отсутствия свойства монотонности у метода индукции над стандартными теориями. Пусть f - немонотонный метод над стандартными эмпирическими теориями. Тогда найдутся тестовые алгоритмы T_0^v, T_1^v, T_2^v и протоколы pr_1^v, pr_2^v (в некотором словаре v), которые удовлетворяют всем условиям свойства монотонности, и найдется протокол $\widetilde{pr}^v$, для которого выполняются равенства $T_1^v(\widetilde{pr}^v) = 0$ и $T_2^v(\widetilde{pr}^v) = 1$. Пусть $h_i = \langle v, obs^v, T_i^v \rangle$ $(i=0,1,2)$ - стандартные теории, $pr_1^v = pr^v(obs^v(A_1), I^{A_1})$ и $pr_2^v = pr^v(obs^v(A_2), I^{A_2})$. Без ограничения общности рассмотрения (ввиду $pr_1^v \subset pr_2^v$ и C_f-4) будем считать, что $A_1 \subset A_2$ и $I^{A_1} = I^{A_2} \restriction A_1$. Отметим, что $T_0^v(\widetilde{pr}^v)=1$ (поскольку $T_2^v(\widetilde{pr}^v) = 1$ и C_f-3).

Рассмотрим эпистемологическую ситуацию G, возникающую после принятия (исследователем) теории $h_1 = f(h_0, pr_1^v, A_1)$, исходя из начальной теории h_0, метода индукции (эпистемологической установки) f и начальных данных pr_1^v (начальной эмпирической информации $obs^v(A_1)$, описываемой протоколом pr_1^v). Ситуация G характеризуется следующим:

I) теорией h_1 данные $\widetilde{pr}^v$ считаются невозможными (ввиду $T_1^v(\widetilde{pr}^v) = 0$);

2) теория h_1 допускает возможность данных pr_2^v (ввиду $T_1^v(pr_2^v) = 1$);

3) если бы в рамках установки f строилась теория, исходя не из данных pr_1^v, а из данных pr_2^v, то в такой теории данные $\widetilde{pr}^v$ считались бы возможными (ввиду $f(h_0, pr_2^v, A_2) = h_2$ и $T_2^v(\widetilde{pr}^v) = 1$);

4) данные pr_2^v представительнее данных pr_1^v (ввиду того, что а) $pr_1^v \subset pr_2^v$ и б) эмпирическая информация $obs^v(A_2)$ включает в себя всю информацию $obs^v(A_1)$; последнее обусловлено стандартностью инструкции obs^v и включением $A_1 \subset A_2$).

Сопоставление фактов I-4 показывает, что ситуация G содержит в себе неопределенность относительно вопроса: возможна или невозможна эмпирическая ситуация, описываемая протоколом $\widetilde{pr}^v$? Действительно, принятая теория h_1 считает данные $\widetilde{pr}^v$ невозможными, но эта же теория допускает возможность данных pr_2^v, индуцирующих возможность $\widetilde{pr}^v$, причем вывод о возможности данных $\widetilde{pr}^v$ более оправдан, так как основан на более представительных данных. Таким образом, в ситуации G оказывается принятой теория (h_1), которая опровергает саму себя, когда она допускает возможность существования другой теории (h_2), опровергающей предсказание принятой. Ясно, что ситуация G методологически неприемлема (заметим, что ситуация G была бы приемлемой, если бы имело место $T_1^v(\widetilde{pr}^v) = T_2^v(\widetilde{pr}^v)$ или $T_1^v(\widetilde{pr}^v) = 1$ & $T_2^v(\widetilde{pr}^v) = 0$; во втором случае принятая теория допускала бы возможность другой, которая уточняла бы предсказание принятой). Поскольку теория h_0, по предположению, методологически корректна и поскольку ситуация G возникает в рамках эпистемологической установки f, то методологическая неприемлемость ситуации G будет указывать на неприемлемость установки f с методологической точки зрения.

4. Покажем, что монотонные методы индукции (над стандартными теориями) существуют. Определим из них два метода: регулярный (по [2]) метод f_1 и S-регулярный (по [4]) метод f_2 (при определении метода f предполагается, что f удовлетворяет исходному свойству C_f-I и явно определяется только отображение ind_f (свойства C_f-2,3,4 метода f будут очевидными следствиями определения ind_f)). Через $\bar{\bar{x}}$ будем обозначать мощность множества x.

Для всякого протокола pr^v в словаре v

$$\mathrm{ind}_{f_1}(T_0^V, pr_0^V)(pr^V) = \begin{cases} I, & \text{если } T_0^V(pr^V)=1 \text{ и } \bar{\bar{B}}(pr^V) < \bar{\bar{B}}(pr_0^V); \\ I, & \text{если } pr^V \simeq pr_0^V; \\ 0 & \text{– во всех остальных случаях;} \end{cases}$$

$$\mathrm{ind}_{f_2}(T_0^V, pr_0^V)(pr^V) = \begin{cases} I, & \text{если } T_0^V(pr^V) = 1,\ \bar{\bar{B}}(pr^V) < \bar{\bar{B}}(pr_0^V) \text{ и существует } D \subset B(pr_0^V) \text{ такое, что } pr^V \simeq pr_0^V \upharpoonright D; \\ I, & \text{если } T_0^V(pr^V) = 1,\ \bar{\bar{B}}(pr^V) \geq \bar{\bar{B}}(pr_0^V) \text{ и для всякого } D \subseteq B(pr^V) \text{ выполняется } \bar{\bar{D}} = \bar{\bar{B}}(pr_0^V) \to pr^V \upharpoonright D \simeq pr_0^V; \\ 0 & \text{– во всех остальных случаях.} \end{cases}$$

Ясно, что методы f_1 и f_2 можно рассматривать в качестве методов над стандартными теориями (о f_2 см. [4, п.8]).

Очевидно, что имеет место

ПРЕДЛОЖЕНИЕ I. Метод индукции f_1 монотонен.

ПРЕДЛОЖЕНИЕ 2. Метод индукции f_2 монотонен.

ДОКАЗАТЕЛЬСТВО. Пусть тестовые алгоритмы T_0^V, T_1^V, T_2^V и протоколы pr_1^V, pr_2^V в словаре v таковы, что $T_0^V(pr_1^V) = T_0^V(pr_2^V) = 1$, $T_1^V = \mathrm{ind}_{f_2}(T_0^V, pr_1^V)$, $T_2^V = \mathrm{ind}_{f_2}(T_0^V, pr_2^V)$, $pr_1^V \subset pr_2^V$ и $T_1^V(pr_2^V) = 1$. И пусть pr^V – протокол в словаре v. Если $T_0^V(pr^V) = 0$, то импликация $T_1^V(pr^V) = 0 \to T_2^V(pr^V) = 0$ очевидна. Покажем, что эта импликация справедлива и при $T_0^V(pr^V) = 1$.

Прежде всего отметим, что (q_1) $\bar{\bar{B}}(pr_1^V) < \bar{\bar{B}}(pr_2^V)$ (следует из $pr_1^V \subset pr_2^V$) и (q_2) для всякого $D \subseteq B(pr_2^V)$ выполняется $\bar{\bar{D}} = \bar{\bar{B}}(pr_1^V) \to pr_2^V \upharpoonright D \simeq pr_1^V$ (следует из $T_1^V(pr_2^V) = 1$ и определения ind_{f_2}).

Возможны три случая.

<u>Случай I.</u> $\bar{\bar{B}}(pr^V) < \bar{\bar{B}}(pr_1^V)$. В этом случае $\bar{\bar{B}}(pr^V) < \bar{\bar{B}}(pr_2^V)$ (ввиду q_1). Пусть $T_2^V(pr^V) = 1$. Поскольку существует множество $D \subset B(pr_2^V)$ такое, что $pr^V \simeq pr_2^V \upharpoonright D$ (по определению ind_{f_2}), и поскольку имеет место q_2, то (ввиду $B(pr^V) < B(pr_1^V) < B(pr_2^V)$) су -

ществует множество $C \subset B(pr_1^V)$ такое, что $pr^V \simeq pr_1^V \restriction C$. Следовательно, $T_1^V(pr^V) = 1$ (по определению ind_{f_2}).

Случай 2. $\bar{\bar{B}}(pr_1^V) \le \bar{\bar{B}}(pr^V) < \bar{\bar{B}}(pr_2^V)$. Пусть $T_1^V(pr^V) = 0$. И пусть множество $D \subseteq B(pr^V)$ таково, что $\bar{\bar{D}} = \bar{\bar{B}}(pr_1^V)$ и $pr^V \restriction D \not\simeq pr_1^V$ (существование такого множества следует из определения ind_{f_2}). Покажем, что равенство $T_2^V(pr^V) = 1$ невозможно. Пусть $T_2^V(pr^V) = 1$. Рассмотрим множество $C \subset B(pr_2^V)$ такое, что $pr^V \simeq pr_2^V \restriction C$ (существование такого множества следует из определения ind_{f_2}). Из $pr^V \simeq pr_2^V \restriction C$ следует существование множества $D_1 \subseteq C$, для которого выполняется $pr^V \restriction D \simeq (pr_2^V \restriction C) \restriction D_1$, т.е. $pr^V \restriction D \simeq pr_2^V \restriction D_1$. Отсюда выводим $pr_1^V \not\simeq pr_2^V \restriction D_1$ (ввиду $pr^V \restriction D \not\simeq pr_1^V$) и $\bar{\bar{D}}_1 = \bar{\bar{D}}$. Учитывая равенство $\bar{\bar{D}} = \bar{\bar{B}}(pr_1^V)$, получаем $\bar{\bar{D}}_1 = \bar{\bar{B}}(pr_1^V)$. Последнее равенство вместе с $pr_1^V \not\simeq pr_2^V \restriction D_1$ приводит к $T_1^V(pr_2^V) = 0$ (ввиду определения ind_{f_2}), что невозможно.

Случай 3. $\bar{\bar{B}}(pr^V) \ge \bar{\bar{B}}(pr_2^V)$. В этом случае $\bar{\bar{B}}(pr^V) \ge \bar{\bar{B}}(pr_1^V)$ (ввиду q_1). Пусть $T_2^V(pr^V) = 1$. Из определения ind_{f_2}, равенств $T_2^V(pr^V) = 1$ и $T_1^V(pr_2^V) = 1$ следует (ввиду $\bar{\bar{B}}(pr^V) \ge \bar{\bar{B}}(pr_2^V)$ и $\bar{\bar{B}}(pr_2^V) \ge \bar{\bar{B}}(pr_1^V)$), что (q_3) для всякого $D \subset B(pr^V)$ выполняется $\bar{\bar{D}} = \bar{\bar{B}}(pr_2^V) \to pr^V \restriction D \simeq pr_2^V$ и (q_4) для всякого $C \subseteq B(pr_2^V)$ выполняется $\bar{\bar{C}} = \bar{\bar{B}}(pr_1^V) \to pr_2^V \restriction C \simeq pr_1^V$. Из (q_3), (q_4) следует, что для всякого $D_1 \subseteq B(pr^V)$ выполняется $\bar{\bar{D}}_1 = \bar{\bar{B}}(pr_1^V) \to pr^V \restriction D_1 \simeq pr_1^V$. Последнее влечет $T_1^V(pr^V) = 1$ (ввиду определения ind_{f_2}).

5. Известно [5], что всякий универсальный алгоритм распознавания (построения решающей функции) является, по сути, некоторым методом индукции, причем методом над стандартными эмпирическими теориями (последнее следует из приведенного ниже определения универсального алгоритма распознавания). В связи с этим представляется полезным исследование (универсальных) алгоритмов распознавания относительно их монотонности. Сформулируем свойство монотонности в терминах, принятых в общей теории распознавания образов.

Пусть $(P_1,\dots,P_N;P_0)$ - набор признаков. Предполагается, что всякому признаку P из этого набора поставлена в соответствие конкретная измерительная процедура, применимая к любому объекту из некоторой (изучаемой) совокупности эмпирических объектов U, и что при измерении признака P у объекта $u\in U$ наблюдается вполне определенное (числовое) значение $P(u)$ этого признака. Множество всех возможных значений признака $P_i, i=1,\dots,N;0$, будем обозначать через Z_i. Признаковым пространством X_0 является непустое подмножество декартова произведения $X=Z_1\times\dots\times Z_N$. Предполагается, что всякий объект $u\in U$ представлен в X_0 точкой $x=\langle P_1(u),\dots$ $\dots,P_N(u)\rangle=\langle p_1(x),\dots,p_N(x)\rangle$, где $p_1,\dots,p_N$ - "признаки" точек из X_0, соответствующие признакам $P_1,\dots,P_N$. Выделенный признак P_0 называется целевым. Множество Z_0 значений этого признака есть множество $\{1,\dots,K\}$, $K\ge 2$, номеров образов. Факт $P_0(u)=j$, $u\in U$, означает, что объект u принадлежит j-му образу. Если $x=$ $=\langle P_1(u),\dots,P_N(u)\rangle$, то $P_0(u)=p_0(x)$, где p_0 - целевой "признак" точек из X_0, соответствующий признаку P_0.

Обучающей выборкой называется пара $\langle w,\varphi\rangle$, где w - конечное собственное подмножество пространства X_0, а φ - отображение множества w на Z_0, интерпретируемое так, что $\forall x\in w(\varphi(x)=p_0(x))$. Если эта обучающая выборка реальная (а не модельная), то всякой точке $x\in w$ соответствует объект $u\in U$ такой, что $\langle P_1(u),\dots$ $\dots,P_N(u)\rangle=x$, значение $P_0(u)$ замерено и $\varphi(x)=P_0(u)$. Решающей функцией называется отображение Ψ: $X_0\to Z_0^0$, где $Z_0^0=Z_0\cup\{0\}$. Предполагается, что если $x\in X_0$ представляет объект $u\in U$, тогда а) если $\Psi(x)=j$, $1\le j\le K$, то предсказывается (решающей функцией Ψ) равенство $P_0(u)=j$ (объект u распознается, хотя, возможно, и неправильно); б) если $\Psi(x)=0$, то не делается никакого предсказания относительно значения $P_0(u)$ (объект u не распознается). Универсальный алгоритм распознавания R по всякой обучающей выборке $W=\langle w,\varphi\rangle$ строит решающую функцию $R(W)$ такую, что а) для всякого $x\in w$ выполняется $R(W)(x)=\varphi(x)$, т.е. $R(W)\restriction w=\varphi$; и б) для всякого $x\notin w$, если $R(W)(x)\ne 0$, то предсказание равенства $p_0(x)=$ $=R(W)(x)$ является наиболее обоснованным из предсказаний $p_0(x)=$ $=1,\dots,p_0(x)=K$.

<u>Свойство монотонности алгоритмов распознавания</u>. Для всяких обучающих выборок $W_1=\langle w_1,\varphi_1\rangle$ и $W_2=\langle w_2,\varphi_2\rangle$, если $w_1\subset w_2$, $\varphi_1=\varphi_2\restriction w_1$ и для любой точки y из дополнения $w_2\setminus w_1$ выполняется $\varphi_2(y)=R(W_1)(y)$, то для всякого $x\in X_0$ имеет место импликация $R(W_1)(x)\ne 0\to R(W_2)(x)=R(W_1)(x)$.

6. Покажем, что существуют как монотонные, так и немонотонные алгоритмы распознавания. Определяемые ниже пять алгоритмов будут иллюстративными в том смысле, что в этих алгоритмах та или иная идея распознавания представлена в "чистом" виде. Такая иллюстративность позволяет теоретически оценивать сами идеи распознавания.

Первые три алгоритма построены методами линейных решающих функций. При формулировке этих алгоритмов предполагается следующее: 1) $N = 2$ и $K = 2$ (ограничения, упрощающие изложение); 2) признаки P_1 и P_2 измеряются в шкале отношений; 3) $Z_1 = Z_2 =$ = множеству всех положительных действительных чисел; 4) X - квадрант евклидовой плоскости с прямоугольной системой координат и 5) $X_0 = X$ (ограничение, упрощающее изложение).

Алгоритм R_1. Входная конъюнкция - обучающая выборка $W = \langle w, \varphi \rangle$. Предполагается, что $w = w^1 \cup w^2$ и для всякой $t \in w$, если $t \in w^1$, то $\varphi(t) = 1$, а если $t \in w^2$, то $\varphi(t) = 2$.

Шаг 1. Вычисляется r_{min} такое, что

а) $\forall t_1 \in w^1 \ \forall t_2 \in w^2 (r(t_1, t_2) \geq r_{min})$,

б) $\exists t_1 \in w^1 \ \exists t_2 \in w^2 (r(t_1, t_2) = r_{min})$.

Комментарий: здесь и далее через $r(t_1, t_2)$ обозначается расстояние между точками t_1 и t_2.

Шаг 2. Из множества $M = \{\langle t_1, t_2 \rangle \mid t_1 \in w^1, t_2 \in w^2, r(t_1, t_2) = r_{min}\}$ выбирается (произвольно) пара $\langle t_0^1, t_0^2 \rangle$.

Комментарий: результат работы алгоритма R_1 не зависит от допускаемого здесь произвола.

Шаг 3. Строятся прямые линии L^1 и L^2 (ограниченные пространством X_0), перпендикулярные отрезку $[t_0^1, t_0^2]$ и такие, что $t_0^1 \in L^1$ и $t_0^2 \in L^2$.

Шаг 4. Направление вдоль отрезка $[t_0^1, t_0^2]$ от точки t_0^1 к точке t_0^2 принимается за направление слева-направо, и строится разбиение $\{X_0^1, X_0^0, X_0^2\}$ пространства X_0 такое, что а) X_0^1 - часть X_0, лежащая левее области X_0^0, б) X_0^0 - часть X_0, лежащая между линиями L^1 и L^2; в) X_0^2 - часть X_0, лежащая правее области X_0^0.

Комментарий: $L^1 \subseteq X_0^1$, $L^2 \subseteq X_0^2$.

Шаг 5. Если $w^1 \not\subseteq X_0^1$ или $w^2 \not\subseteq X_0^2$, то перейти на шаг 7.

Шаг 6. Принимается, что для всякой $t \in X_0$

$$R_1(W)(t) = \begin{cases} 1, & \text{если } t \in X_0^1; \\ 2, & \text{если } t \in X_0^2; \\ 0, & \text{если } t \in X_0^0; \end{cases}$$

перейти на шаг 8.

Шаг 7. Принимается, что для всякой $t \in X_0$

$$R_1(W)(t) = \begin{cases} \varphi(t), & \text{если } t \in w, \\ 0, & \text{если } t \notin w. \end{cases}$$

Шаг 8. Конец.

Алгоритм R_2. Описание алгоритма R_2 совпадает с описанием алгоритма R_1 (при замене, естественно, выражения "R_1" на "R_2" во всяком вхождении) за исключением шага 6, который здесь формулируется следующим образом:

Шаг 6. Строится прямая линия L^0 (ограниченная пространством X_0), перпендикулярная отрезку $[t_0^1, t_0^2]$ и проходящая через середину этого отрезка, и принимается, что для всякой $t \in X_0$

$$R_2(W)(t) = \begin{cases} 1, & \text{если } t \text{ расположена левее линии } L^0; \\ 2, & \text{если } t \text{ расположена правее линии } L^0; \\ 0, & \text{если } t \in L^0; \end{cases}$$

перейти на шаг 8.

Алгоритм R_3. Входная информация для R_3 совпадает с входной информацией для R_1 .

Шаг 1. Если мощность $\bar{\bar{w}}^1 < 3$ или $\bar{\bar{w}}^2 < 3$, то перейти на шаг 6.

Шаг 2. Вычисляется i

$$i := \begin{cases} 1, & \text{если } \exists L_1(w^1 \subseteq L_1) \mathbin{\&} \neg\exists L_2(w^2 \subseteq L_2); \\ 2, & \text{если } \exists L_2(w^2 \subseteq L_2) \mathbin{\&} \neg\exists L_1(w^1 \subseteq L_1); \\ 0 & \text{– во всех остальных случаях.} \end{cases}$$

Комментарий: областью значений переменных L_1 и L_2 является множество всех прямых линий (ограниченных пространством X_0).

Шаг 3. Если $i = 0$, то перейти на шаг 6.

Шаг 4. Строится прямая линия L^i такая, что $w^i \subseteq L^i$.

Шаг 5. Принимается, что для всякой $t \in X_0$

$$R_3(W)(t) = \begin{cases} i, \text{ если } t \in L^i; \\ I, \text{ если } t \notin L^i \text{ и } i = 2; \\ 2, \text{ если } t \notin L^i \text{ и } i = 1; \end{cases}$$

перейти на шаг 7.

Шаг 6. Принимается, что для всякой $t \in X_0$

$$R_3(W)(t) = \begin{cases} \varphi(t), \text{ если } t \in w, \\ 0, \text{ если } t \notin w. \end{cases}$$

Шаг 7. Конец.

ПРЕДЛОЖЕНИЕ 3. *Алгоритм распознавания R_1 монотонен.*

ДОКАЗАТЕЛЬСТВО. Пусть $W = \langle w, \varphi \rangle$ и $\widetilde{W} = \langle \widetilde{w}, \widetilde{\varphi} \rangle$ - обучающие выборки для алгоритма R_1 и пусть $w \subset \widetilde{w}$, $\varphi = \widetilde{\varphi} \restriction w$ и $\forall t \in \widetilde{w} \setminus w (\widetilde{\varphi}(t) = R_1(W)(t))$. Покажем, что $R_1(\widetilde{W}) = R_1(W)$. Всякое используемое ниже непоясненное обозначение, которое не содержит волну "$\sim$", будет взято из описания алгоритма R_1 (обрабатывающего выборку W). Обозначаемое через "$\widetilde{S}$" будет аналогично обозначаемому через "S" и связано либо с выборкой $\widetilde{W}$, либо с обработкой этой выборки алгоритмом R_1.

Ясно, что функция $R_1(w)$ не получается на шаге 7 (алгоритма R_1), поскольку $w \subset \widetilde{w}$, $\varphi = \widetilde{\varphi} \restriction w$ и $\widetilde{\varphi}$ - отображение на $\{1,2\}$. Следовательно, $R_1(w)$ получается на шаге 6 и $w^1 \subseteq X_0^1$ & $w^2 \subseteq X_0^2$.Очевидно, что для доказательства равенства $R_1(\widetilde{W}) = R_1(W)$ достаточно доказать справедливость утверждения (q) $\widetilde{L}^1 = L^1$ & $\widetilde{L}^2 = L^2$& $\widetilde{w}^1 \subseteq \widetilde{X}_0^1$ & $\widetilde{w}^2 \subseteq \widetilde{X}_0^2$.В свою очередь,для доказательства (q) достаточно показать,что (q) имеет место во всех ситуациях, когда мощность множества $\widetilde{w}$ больше мощности множества w точно на единицу.

Пусть $\widetilde{w} = w \cup \{t^*\}$, $t^* \notin w$. Из $R_1(W)(t^*) \neq 0$ следует,что либо $t^* \in X_0^1$, либо $t^* \in X_0^2$. Поскольку эти альтернативы логически симметричны, рассмотрим только одну из них. Пусть $t^* \in X_0^1$.Тогда $\widetilde{w}^1 = w^1 \cup \{ t^* \}$ и $\widetilde{w}^2 = w^2$. Ясно, что $\forall t \in \widetilde{w}^2 (r(t^*,t) \geq r_{min})$,поскольку $t^* \in X_0^1$, а всякая t из $\widetilde{w}^2 = w^2$ является точкой из X_0^2.

Следовательно, $\tilde{r}_{min} = r_{min}$ и $\langle t_0^1, t_0^2 \rangle \in \tilde{M}$. Значит, $\tilde{L}^1 = L^1$ и $\tilde{L}^2 = L^2$. Далее, $\tilde{w}^1 \subseteq X_0^1 = \tilde{X}_0^1$, и $\tilde{w}^2 \subseteq X_0^2 = \tilde{X}_0^2$. Таким образом, имеет место (q).

ПРЕДЛОЖЕНИЕ 4. Алгоритм распознавания R_2 немонотонен.

ДОКАЗАТЕЛЬСТВО. Зафиксируем в X_0 четыре точки: $t_1 = \langle 1,1 \rangle$, $t_2 = \langle 1,5 \rangle$, $t_3 = \langle 1,7 \rangle$ и $t_4 = \langle 1,11 \rangle$. Пусть для R_2 обучающая выборка $W_1 = \langle w_1, \varphi_1 \rangle$, где $w_1 = \{t_1, t_4\}$, $\varphi_1(t_1) = 1$ и $\varphi_1(t_4) = 2$, а выборка $W_2 = \langle w_2, \varphi_2 \rangle$, где $w_2 = \{t_1, t_3, t_4\}$, $\varphi_2(t_1) = I$ и $\varphi_2(t_3) = \varphi_2(t_4) = 2$. Непосредственной проверкой вычислений, осуществляемых алгоритмом R_2, устанавливается следующее: $R_2(W_1)(t_2) = 1$, $R_2(W_1)(t_3) = 2$, $R_2(W_2)(t_2) = 2$. Ясно, что $w_1 \subset w_2$, $\varphi_1 = \varphi_2 \restriction w_1$ и для всякой $t \in w_2 \setminus w_1$ выполняется $\varphi_2(t) = R_2(W_1)(t)$. Однако $R_2(W_1)(t_2) \neq 0$ и $R_2(W_2)(t_2) \neq R_2(W_1)(t_2)$.

ПРЕДЛОЖЕНИЕ 5. Алгоритм распознавания R_3 монотонен.

Доказательство легко усматривается из описания алгоритма.

Четвертый алгоритм R_4 построен в рамках методики логических решающих функций [6], а пятый R_5 – в рамках методики эмпирического предсказания, описанной в [7]. При формулировке этих алгоритмов предполагается следующее: I) $N = 3$ и $K = 2$; 2) признаки P_1, P_2 и P_3 измеряются в шкале наименований; 3) $Z_1 = \{0,1\}$ для алгоритма R_4 и $Z_1 = \{0,1,2\}$ для алгоритма R_5, $Z_2 = Z_3 = \{0,1,2\}$ и 4) $X_0 = X$ (ограничения I–4 упрощают изложение). Через S_4 будем обозначать множество $\{s_1(x), \ldots, s_7(x)\}$ (элементарных) свойств точек пространства X_0, содержащее свойства $p_1(x) = 0$ и $p_i(x) = j$ для $i = 2,3$ и $j = 0,I,2$. Через S_5 будем обозначать множество $\{p_i(x) = j\}$, $i = I,2,3$, $j = 0,I,2$. Формулу вида $\forall x \in X_0 (a(x) \to c(x))$ будем называть закономерностью (над X_0), если $a(x)$ – конъюнкция, элементами которой являются свойства из S_5 или их отрицания, а $c(x)$ есть $p_0(x) = 1$ или $p_0(x) = 2$ (заметим, что $p_0(x) \neq 1 \leftrightarrow p_0(x) = 2$).

Алгоритм R_4. Входная информация – обучающая выборка $W = \langle w, \varphi \rangle$. Комментарий: основой алгоритма R_4 является построение множества M конъюнкций (сложных свойств точек из X_0); элементами конъюнкций из M могут быть только свойства из S_4 или их отрицания.

Шаг I. $M := \emptyset$; $V := w$; $S := S_4$; перейти на шаг 5.

Шаг 2. Если $M = \emptyset$, то перейти на шаг 9.

Шаг 3. В множестве M ищется конъюнкция $q(x)$ такая, что

$$\exists t_1, t_2 \in w(q(t_1) \,\&\, q(t_2) \,\&\, \varphi(t_1) \neq \varphi(t_2)).$$

Если такая конъюнкция не находится, то перейти на шаг 8.

Комментарий: если искомых конъюнкций несколько, то $q(x)$ - выбранная из них произвольно; результат работы алгоритма R_4 не зависит от этого произвола.

Шаг 4. $V := \{t \in w \mid q(t)\}$; $S := \{s(x) \in S_4 \mid \exists t_1, t_2 \in V(s(t_1) \,\&\, \neg s(t_2))\}$.
Комментарий: $S \neq \emptyset$.

Шаг 5. Для каждого свойства $s_i(x) \in S$ вычисляются числа:

m_i^1 - число точек в множестве $\{t \in V \mid s_i(t) \,\&\, \varphi(t)=1\}$;

m_i^2 - число точек в множестве $\{t \in V \mid s_i(t) \,\&\, \varphi(t)=2\}$;

n_i^1 - число точек в множестве $\{t \in V \mid \neg s_i(t) \,\&\, \varphi(t)=1\}$;

n_i^2 - число точек в множестве $\{t \in V \mid \neg s_i(t) \,\&\, \varphi(t)=2\}$;

$$l_i = \max(m_i^1, m_i^2) + \max(n_i^1, n_i^2) \;.$$

Шаг 6. В множестве S ищется свойство $s_k(x)$ такое, что для всякого $s_j(x) \in S$, отличного от $s_k(x)$, выполняется $l_j < l_k$; если такое свойство не находится, тогда [[если $M \neq \emptyset$, то из M удаляется конъюнкция $q(x)$] ; перейти на шаг 2].

Шаг 7. Если $M = \emptyset$, то в M вносятся конъюнкции (свойства) $s_k(x)$ и $\neg s_k(x)$, иначе [из M удаляется конъюнкция $q(x)$; в M вносятся конъюнкции $q(x) \,\&\, s_k(x)$ и $q(x) \,\&\, \neg s_k(x)$]; перейти на шаг 3.

Шаг 8. Принимается следующее: для всякой $t \in X_0$

$$R_4(W)(t) = \begin{cases} 1, & \text{если } \exists q(x) \in M(q(t) \,\&\, \exists t_1 \in w(q(t_1) \,\&\, \varphi(t_1)=1)); \\ 2, & \text{если } \exists q(x) \in M(q(t) \,\&\, \exists t_1 \in w(q(t_1) \,\&\, \varphi(t_1)=2)); \\ \varphi(t), & \text{если } \forall q(x) \in M(\neg q(t)) \,\&\, t \in w; \\ 0, & \text{если } \forall q(x) \in M(\neg q(t)) \,\&\, t \notin w; \end{cases}$$

перейти на шаг 10.

Шаг 9. Принимается следующее: для всякой $t \in X_0$

$$R_4(W)(t) = \begin{cases} \varphi(t), & \text{если } t \in w \;, \\ 0, & \text{если } t \notin w \;. \end{cases}$$

Шаг 10. Конец.

Алгоритм R_5. Входная конъюнкция – обучающая выборка $W=\langle w,\varphi\rangle$. Априорная информация (параметр алгоритма): множество закономерностей $Q_0=\{z_i\}_{i=1,\dots,k}$ такое, что для всякого $i=1,\dots,k$ свойство $c_i(x)$ есть $p_0(x)=1$. Комментарий: всякая $z_i\in Q_0$ есть $\forall x\in X_0(a_i(x)\to c_i(x))$.

Шаг 1. Для каждой закономерности $z_i\in Q_0$ вычисляются числа:

m_i – число точек в множестве $\{t\in w \mid a_i(t)\ \&\ \varphi(t)=1\}$;

n_i – число точек в множестве $\{t\in w \mid a_i(t)\}$;

$$l_i=\begin{cases}\dfrac{m_i}{n_i}, & \text{если } n_i\neq 0;\\[2mm] \dfrac{1}{2}, & \text{если } n_i=0.\end{cases}$$

Шаг 2. Принимается следующее: для всякой $t\in X_0$ значение $R_5(W)(t)$ определяется подалгоритмом, состоящим из шагов 2_1–2_7.

Шаг 2_1. Если $t\in w$, то [$R_5(W)(t):=\varphi(t)$; перейти на шаг 2_7].

Шаг 2_2. Строится множество $C=\{\langle\rho_i,\mu_i\rangle\}_{i=1,\dots,k}$ пар такое, что для $i=1,\dots,k$

$$\langle\rho_i,\mu_i\rangle=\begin{cases}\langle 1,l_i\rangle, & \text{если } a_i(t)\ \&\ l_i>\frac{1}{2};\\ \langle -1,1-l_i\rangle, & \text{если } a_i(t)\ \&\ l_i<\frac{1}{2};\\ \langle 0,0\rangle & \text{– в остальных случаях.}\end{cases}$$

Комментарий: вхождение в C пары $\langle 1,\mu_i\rangle$ ($\langle -1,\mu_i\rangle$) означает, что закономерность z_i "голосует" за $p_0(t)=1$, ($p_0(t)=2$) с уверенностью μ_i.

Шаг 2_3. Если $\exists i(\langle\rho_i,1\rangle\in C)$, то

$$[R_5(W)(t):=\begin{cases}1, & \text{если } \forall j(\langle\rho_j,1\rangle\in C\to\rho_j=1);\\ 2, & \text{если } \forall j(\langle\rho_j,1\rangle\in C\to\rho_j=-1);\\ 0 & \text{– в противном случае.}\end{cases}$$

перейти на шаг 2_7].

Шаг 2_4. Цикл по $i=1,\dots,k-1$; цикл по $j=i+1,\dots,k$: если $\rho_i\neq\rho_j\ \&\ \mu_i\neq 0\ \&\ \mu_j\neq 0$, то [формируется закономерность $z_{ij}=$

$= \forall x \in X_0(a_i(x) \,\&\, a_i(x) \to p_0(x) = 1)$; аналогично тому, как это делается на шаге I, вычисляется l_{ij} ; если $l_{ij} > \frac{1}{2}$ $(l_{ij} < \frac{1}{2})$, то заменяется на $\langle 0, 0\rangle$ та пара из $\{\langle \rho_i, \mu_i\rangle, \langle \rho_j, \mu_j\rangle\}$, первый элемент которой равен -1 (1)].

Шаг 2_5. $\mu_{max} := \max(\mu_1, \ldots, \mu_k)$.

Шаг 2_6.

$$R_5(W)(t) := \begin{cases} 1, \text{ если } \langle 1, \mu_{max}\rangle \in C \,\&\, \langle -1, \mu_{max}\rangle \notin C; \\ 2, \text{ если } \langle -1, \mu_{max}\rangle \in C \,\&\, \langle 1, \mu_{max}\rangle \notin C; \\ 0 - \text{в остальных случаях.} \end{cases}$$

Шаг 2_7. Конец.

Шаг 3. Конец.

ПРЕДЛОЖЕНИЕ 6. Алгоритм распознавания R_4 немонотонен.

ДОКАЗАТЕЛЬСТВО. Зафиксируем в X_0 II точек: $t_1 = \langle 0,2,0\rangle$, $t_2 = \langle 1,0,1\rangle$, $t_3 = \langle 1,1,2\rangle$, $t_4 = \langle 0,1,1\rangle$, $t_5 = \langle 0,2,2\rangle$, $t_6 = \langle 1,0,0\rangle$, $t_7 = \langle 0,1,0\rangle$, $t_8 = \langle 1,1,1\rangle$, $t_9 = \langle 0,0,2\rangle$, $t_{10} = \langle 1,2,0\rangle$, $t_{11} = \langle 0,0,1\rangle$. Пусть для R_4 обучающая выборка $W_1 = \langle w_1, \varphi_1\rangle$, где $w_1 = \{t_1, \ldots, t_6\}$, $\varphi_1(t_1) = \varphi_1(t_2) = \varphi_1(t_3) = 1$ и $\varphi_1(t_4) = \varphi_1(t_5) = \varphi_1(t_6) = 2$, а выборка $W_2 = \langle w_2, \varphi_2\rangle$, где $w_2 = \{t_1, \ldots, t_{10}\}$, $\varphi_2 \upharpoonright w_1 = \varphi_1$, $\varphi_2(t_7) = \varphi_2(t_8) = 1$ и $\varphi_2(t_9) = \varphi_2(t_{10}) = 2$. Непосредственной проверкой вычислений, осуществляемых алгоритмом R_4, устанавливается, что для всякой $t \in X_0$

$$R_4(W_1)(t) = \begin{cases} 1, \text{ если } p_1(t)=0 \,\&\, p_3(t)=0 \text{ или } p_1(t)\neq 0 \,\&\, p_3(t)\neq 0; \\ 2, \text{ если } p_1(t)=0 \,\&\, p_3(t)\neq 0 \text{ или } p_1(t)\neq 0 \,\&\, p_3(t)=0; \end{cases}$$

$$R_4(W_2)(t) = \begin{cases} 1, \text{ если } p_2(t)\neq 1 \,\&\, p_3(t)=1; \\ \varphi_2(t), \text{ если } (p_2(t)=1 \vee p_3(t)\neq 1) \,\&\, t \in w_2; \\ 0, \text{ если } (p_2(t)=1 \vee p_3(t)\neq 1) \,\&\, t \notin w_2. \end{cases}$$

Ясно, что $w_1 \subset w_2$, $\varphi_1 = \varphi_2 \upharpoonright w_1$ и для всякой $t \in w_2 \setminus w_1$ выполняется $\varphi_2(t) = R_4(W_1)(t)$. Однако $R_4(W_1)(t_{11}) \neq 0$ и $R_4(W_2)(t_{11}) \neq R_4(W_1)(t_{11})$.

ПРЕДЛОЖЕНИЕ 7. Алгоритм распознавания R_5 немонотонен.

ДОКАЗАТЕЛЬСТВО. Пусть параметр $Q_0 = \{z_1, z_2\}$, где z_1 есть $\forall x \in X_0 (p_1(x) = 0 \to p_0(x) = 1)$, а z_2 есть $\forall x \in X_0(p_2(x)=0 \to p_0(x) = 1)$. Зафиксируем в X_0 10 точек: $t_1 = \langle 0,1,0 \rangle$, $t_2 = \langle 0,1,1 \rangle$, $t_3 = \langle 1,0,0 \rangle$, $t_4 = \langle 0,1,2 \rangle$, $t_5 = \langle 1,0,1 \rangle$, $t_6 = \langle 1,0,2 \rangle$, $t_7 = \langle 2,0,0 \rangle$, $t_8 = \langle 0,2,0 \rangle$, $t_9 = \langle 0,2,1 \rangle$, $t_{10} = \langle 0,0,0 \rangle$. Пусть для R_5 обучающая выборка $W_1 = \langle w_1, \varphi_1 \rangle$, где $w_1 = \{t_1, \ldots, t_7\}$, $\varphi_1(t_1) = \varphi_1(t_2) = \varphi_1(t_3) = 1$ и $\varphi_1(t_4) = \ldots = \varphi_1(t_7) = 2$, а выборка $W_2 = \langle w_2, \varphi_2 \rangle$, где $w_2 = \{t_1, \ldots, t_9\}$, $\varphi_2 \restriction w_1 = \varphi_1$ и $\varphi_2(t_8) = \varphi_2(t_9) = 1$. Непосредственной проверкой вычислений, осуществляемых алгоритмом R_5, устанавливается следующее: $R_5(W_1)(t_8) = R_5(W_1)(t_9) = 1$, $R_5(W_1)(t_{10}) = 2$ и $R_5(W_2)(t_{10}) = 1$. Ясно, что $w_1 \subset w_2$, $\varphi_1 = \varphi_2 \restriction w_1$ и для всякой $t \in w_2 \setminus w_1$ выполняется $\varphi_2(t) = R_5(W_1)(t)$. Однако $R_5(W_1)(t_{10}) \neq 0$ и $R_5(W_2)(t_{10}) \neq R_5(W_1)(t_{10})$.

7. Метод индукции f будем называть m-неустойчивым, $m > 1$, тогда и только тогда, когда существуют тестовые алгоритмы T_0^v, $T_1^v, \ldots, T_m^v$ в (некотором) словаре v и протоколы $pr_1^v, \ldots, pr_m^v$ в словаре v такие, что а) для всякого $i = 1, \ldots, m$ выполняются $T_0^v(pr_i^v) = 1$ и $T_i^v = \mathrm{ind}_f(T_0^v, pr_i^v)$; б) для всякого $i = 2, \ldots, m$ выполняются $pr_{i-1}^v \subset pr_i^v$ и $T_1^v(pr_i^v) = 1$ и в) существует протокол $\widetilde{pr}^v$ в словаре v, для которого имеет место

$$T_1^v(\widetilde{pr}^v) = 0 \ \& \ \forall i \in \{2, \ldots, m\}(T_i^v(\widetilde{pr}^v) \neq T_{i-1}^v(\widetilde{pr}^v)) .$$

Метод индукции f будем называть устойчивым, если и только если метод f не является m-неустойчивым для любого $m > 1$.

ПРЕДЛОЖЕНИЕ 8. *Всякий монотонный метод индукции устойчив.*

Доказательство очевидно.

8. Алгоритм распознавания R будем называть m-неустойчивым, $m > 1$, тогда и только тогда, когда существуют обучающие выборки $W_1 = \langle w_1, \varphi_1 \rangle, \ldots, W_m = \langle w_m, \varphi_m \rangle$ такие, что а) для всякого $i = 2, \ldots, m$ выполняются $w_{i-1} \subset w_i$, $\varphi_1 = \varphi_i \restriction w_1$ и $\forall y \in w_i \setminus w_1 (\varphi_i(y) = R(W_1)(y))$ и б) существует точка $x \in X_0$, для которой имеет место

$$\forall i \in \{1, \ldots, m\}(R(W_i)(x) \neq 0) \ \& \ \forall j \in \{2, \ldots, m\}(R(W_j)(x) \neq R(W_{j-1})(x)) .$$

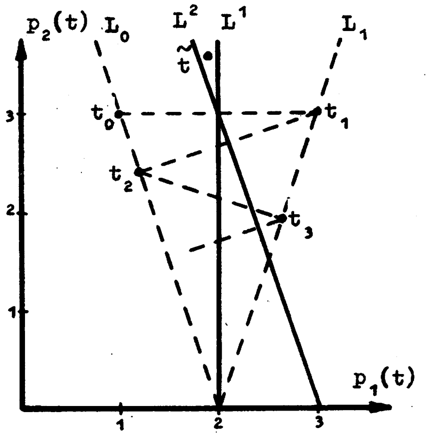
Алгоритм распознавания R будем называть устойчивым, если и только если алгоритм R не является m-неустойчивым для любого $m > 1$.

ПРЕДЛОЖЕНИЕ 9. Всякий монотонный алгоритм распознавания устойчив.

Доказательство очевидно.

ПРЕДЛОЖЕНИЕ IO. Алгоритм распознавания R_2 m-неустойчив для любого $m>1$.

ДОКАЗАТЕЛЬСТВО. На евклидовой плоскости зафиксируем три вспомогательные прямые L_0, L^1 и L_1 (см. рисунок), проходящие через точку $\langle 2,0 \rangle$ и соответственно через точки $\langle 1,3 \rangle$, $\langle 2,3 \rangle$ и $\langle 3,3 \rangle$. Зафиксируем бесконечную последовательность точек t_0, $t_1,\dots$ такую, что $t_0 = \langle 1,3 \rangle$, $t_1 = \langle 3,3 \rangle$ и для каждого $i>1$ точка t_i есть точка пересечения линии L_j, где $j \in \{0,1\}$ удовлетворяет сравнению $i \equiv j(\mathrm{mod}\,2)$, и перпендикуляра, опущенного на эту линию из точки t_{i-1}. Зафиксируем вспомогательную прямую L^2, перпендикулярную отрезку $[t_1,t_2]$ и проходящую через середину этого отрезка. Зафиксируем точку $\tilde{t}$, лежащую в секторе, образованном линиями L^1 и L^2.



Пусть m – натуральное число, большее I. Рассмотрим последовательность обучающих выборок $W_1 = \langle w_1,\varphi_1 \rangle,\dots, W_m = \langle w_m,\varphi_m \rangle$ таких, что для всякого $i=1,\dots,m$ множество w_i есть $\{t_0,t_1,\dots$ $\dots,t_i\}$, а отображение $\varphi_i\colon\ w_i \to \{1,2\}$ удовлетворяет условию

$$\varphi_i(t) = \begin{cases} 1, \text{ если } & t \in L_0, \\ 2, \text{ если } & t \in L_1. \end{cases}$$

Ясно, что для любого $i = 2,\dots,m$ выполняются $w_{i-1} \subset w_i$ и $\varphi_1 = \varphi_i \upharpoonright w_1$. Непосредственной проверкой вычислений, осуществляемых алгоритмом R_2, устанавливается, что а) для всякой $t \in X_0$

$$R_2(W_1)(t) = \begin{cases} 1, & \text{если } p_1(t) < 2, \\ 2, & \text{если } p_1(t) > 2, \\ 0, & \text{если } p_1(t) = 2 \end{cases}$$

(следовательно, для любого $i=2,\ldots,m$ имеет место $\forall t \in w_i \setminus w_1 (\varphi_i(t) = R_2(W_1)(t))$), и б) для всякого $i = 1,\ldots, m$

$$R_2(W_i)(\tilde{t}) = \begin{cases} 1, & \text{если } i \text{ нечетно}; \\ 2, & \text{если } i \text{ четно}. \end{cases}$$

ПРЕДЛОЖЕНИЕ II. Алгоритм распознавания R_5 m-неустойчив для $m = 3$.

ДОКАЗАТЕЛЬСТВО. Пусть параметр Q_0 и обучающие выборки W_1, W_2 – те же, что и в доказательстве предложения 7. Зафиксируем дополнительно в X_0 две точки: $t_{11} = \langle 2,0,1 \rangle$ и $t_{12} = \langle 2,0,2 \rangle$. Пусть обучающая выборка $W_3 = \langle w_3, \varphi_3 \rangle$, где $w_3 = w_2 \cup \{t_{11}, t_{12}\}$, $\varphi_3 \restriction w_2 = \varphi_2$ и $\varphi_3(t_{11}) = \varphi_3(t_{12}) = 2$. Непосредственной проверкой вычислений, осуществляемых алгоритмом R_5, устанавливается (дополнительно к тому, что было установлено в доказательстве предложения 7) следующее: $R_5(W_1)(t_{11}) = R_5(W_1)(t_{12}) = 2$, $R_5(W_3)(t_{10}) = 2$. Ясно, что $w_1 \subset w_2 \subset w_3$, $\varphi_1 = \varphi_2 \restriction w_1$, $\varphi_1 = \varphi_3 \restriction w_1$ и для всякой $t \in w_i \setminus w_1$, $i = 2,3$, выполняется $\varphi_i(t) = R_5(W_1)(t)$. Однако $R_5(W_i)(t_{10}) \neq 0$, $i = 1,2,3$, и $R_5(W_1)(t_{10}) \neq R_5(W_2)(t_{10}) \neq R_5(W_3)(t_{10})$.

Заметим, что доказательство m-неустойчивости алгоритма R_5 для $m > 3$, по-видимому, невозможно, но эта невозможность обусловлена только малостью объема (мощности) признакового пространства X_0.

9. Свойство монотонности у методов индукции над стандартными эмпирическими теориями (следовательно, и у алгоритмов распознавания) представляется необходимым, поскольку немонотонность метода индукции свидетельствует о факте ничем не оправданного произвола, допускаемого в рамках этого метода при построении новых (усиленных) тестовых алгоритмов (новых теорий). Предложение 10 показывает, что немонотонный метод индукции может обладать весьма нежелательным свойством – может быть m-неустойчивым для достаточно больших m.

Литература

1. САМОХВАЛОВ К.Ф. О теории эмпирических предсказаний.–В кн.: Вычислительные системы, вып. 55. Новосибирск, 1973, с. 3–35.

2. ЗАГОРУЙКО Н.Г., САМОХВАЛОВ К.Ф., СВИРИДЕНКО Д.И. Логика эмпирических исследований. -Новосибирск, 1978. - 65 с.

3. НУДЕЛЬМАН А.С. Об одной постановке проблемы индукции. -В кн.: Обнаружение эмпирических закономерностей с помощью ЭВМ (Вычислительные системы, вып. 102). Новосибирск, 1984, с. 54-66.

4. НУДЕЛЬМАН А.С. Об одном методе индукции над стандартными эмпирическими теориями. -В кн.: Методы анализа данных (Вычислительные системы, вып. 111). Новосибирск, 1985, с. 128-139.

5. ЗАГОРУЙКО Н.Г., САМОХВАЛОВ К.Ф. Природа проблемы распознавания образов. -В кн.: Вычислительные системы, вып. 36. Новосибирск, 1969. с.3-12.

6. ЛБОВ Г.С. Методы обработки разнотипных экспериментальных данных. -Новосибирск: Наука, 1981. - 160 с.

7. ВИТЯЕВ Е.Е. Алгоритм эмпирического предсказания. -В кн.: Вычислительные системы, вып. 61. Эмпирическое предсказание и распознавание образов. Новосибирск, 1975, с. 28-36.

Поступила в ред.-изд.отд.
24 января 1986 года