© 2019 **Ю.В. Нагребецкая**[1], канд. физ.-мат. наук, **В.Г. Панов**[2], канд. физ.-мат. наук
[1] *Уральский федеральный университет, Екатеринбург,*
[2] *Институт промышленной экологии УрО РАН, Екатеринбург*

СТЕПЕНЬ ВЗАИМОДЕЙСТВИЯ БИНАРНЫХ ФАКТОРОВ В ТЕОРИИ ДОСТАТОЧНЫХ ПРИЧИН

В рамках формальной модели бинарной теории достаточных причин, основанной на теории конечных булевых алгебр, рассматривается задача определения наличия и «силы» взаимодействия бинарных факторов в данном отклике. Вводятся целочисленные характеристики, которые позволяют классифицировать все отклики в данном бинарном опыте с точки зрения степени присутствия в нем взаимодействия.

Ключевые слова: теория достаточных причин, булевы алгебры, булевы функции, булев куб, отклик, причинность в эпидемиологии, действие группы на множестве.

© 2019 **J.V. Nagrebetskaya**, **V.G. Panov**
Ural Federal UniversityInstitute of Industrial Ecology, Ural Branch of RAS, Ekaterinburg

A DEGREE OF INTERACTION OF BINARY FACTORS IN SUFFICIENT CAUSE THEORY

The problem of indirect determination of similar mechanisms of toxic effects of multifactorial exposure of toxic substances is considered. It is shown that some combinatorial and probabilistic statements allow us to strictly formulate and test a hypothesis of the existence of identical mechanisms of toxic effects for different sets of multifactorial exposure components.

Key words: sufficient cause theory, Boolean algebra, Boolean functions, Boolean cube, outcome, causality in epidemiology, group action on a set.

Введение. Одна из важнейших задач современной эпидемиологии, медицины и экологии состоит в определении причин того или иного наблюдаемого эффекта. Однако понятие *причины* является, с одной стороны, интуитивно ясным, с другой, очень трудно формализуемым, а это последнее совершенно необходимо для использования математических методов. Одной из немногих концепций причинности, которую удается достаточно точно формализовать, является теория достаточных причин, в современном виде представленная в работах философов [1, 2] и эпидемиологов [3, 4].

Фактически необходимость такого формализованного рассмотрения возникает только в эпидемиологии и связанных с ней токсикологии, фармакологии и других медицинских науках, при этом используются модели только для двухуровневых факторов и такого же отклика (бинарная теория) [5–7].

Современное видение формализованной теории достаточных причин представлено в работах [6, 7], которые содержат основные конструкции, позволяющие построить такую модель для n бинарных факторов с подразумеваемыми, но не формулируемыми явно симметриями. Как показано в работах [8–11], адекватная формализация теории достаточных причин может быть сделана на основе теории булевых алгебр. Кроме вопросов формализации теории в работах [5–11] также рассматривается важнейшая проблема определения того, представляет ли данный отклик, понимаемый в контексте теории достаточных причин, взаимодействие действующих n (бинарных) факторов.

Строгая постановка проблемы классификации откликов с точки зрения наличия какого-либо типа совместного действия независимых факторов в этих откликах возможна также на основе теории булевых алгебр. В этом случае булеву алгебру необходимо рассматривать вместе с некоторой группой действующих на ней автоморфизмов [8–11]. Это позволяет провести такую классификацию путем вычисления орбит действия группы автоморфизмов на булевой алгебре откликов. Вычисления показывают, что такая алгебраическая классификация несколько уточняет классификацию, известную для двух факторов и полученную с помощью неформальных рассуждений [6, 8].

Однако, как видно из представленных в работах [9–11] результатов, даже в случае трех действующих бинарных факторов классов различных типов совместного действия оказывается намного больше, чем для двух факторов (22 против 6) и, главное, структура этих классов такова, что какая-либо их предметная трактовка по аналогии с двухфакторным случаем не представляется возможной.

Постановка задачи. Пусть в рамках формальной модели бинарной теории достаточных причин [8–11] получено разбиение на типы взаимодействия всего множества откликов в данном эксперименте. Требуется более формально определить понятие совместного действия (взаимодействия) факторов и ввести такую целочисленную характеристику каждого отклика, которая была бы инвариантна относительно действия группы симметрий данного эксперимента (таким образом, фактически была бы определена на классах различных типов совместного действия) и выражала бы в определённом смысле силу совместного действия факторов в этом отклике.

Формализация задачи. Ниже будут использоваться понятия и определения, приведенные в работах [8–11] и основанные на следующем сопоставлении: действующие бинарные факторы представляются булевыми переменными $x_1,\ldots,x_n$, а отклик представляется булевой функцией $f\left(x_1,\ldots,x_n\right)$ от этих переменных. Через x^{α} будем обозначать конъюнкцию $x_1^{\alpha_1}\ldots x_n^{\alpha_n}, \alpha_i \in \{0,1\}, x^0 = \overline{x}, x^1 = x.$ Булев куб размерности *n* будет обозначаться, как обычно, B^n. Так как, строго говоря, типологизация совместного действия факторов зависит от группы симметрий эксперимента [8–11], то ниже мы ограничимся только случаем классической теории достаточных причин в эпидемиологии, для которой группа симметрий есть группа всех симметрий булева куба B^n. Соответствующий этой группе тип совместного действия с представителем f будем обозначать $\{f\}$. По поводу общей информации о булевых алгебрах и дополнительных понятиях см., например, [12].

Основные результаты. ***Определение*** 1. Будем говорить, что между факторами $x_1,\ldots,x_n$ в отклике *f имеется взаимодействие*, если для некоторого $\alpha \in \mathrm{B}^n$ конъюнкция x^{α} является простой импликантой булевой функции *f*. В этом случае будем говорить, что в отклике *f* взаимодействие достигается при уровнях факторов $x = \alpha$. Это определение вполне согласуется с конструкциями из работ [6, 7].

Нетрудно проверить, что свойство наличия взаимодействия факторов выполняется или не выполняется одновременно для всех булевых функций данного класса $\{f\}$, т.е. является свойством класса. Таким образом, можно говорить о *классах (типах) взаимодействия*.

Теорема 1. Для $n=2$ классами, представляющими взаимодействие *n* факторов, являются только следующие: $\{x_1x_2\}, \{x_1x_2 \vee \overline{x}_1\overline{x}_2\}$. Для $n=3$ такими классами являются только следующие:

$$\{x_1x_2x_3\}, \{x_1x_2x_3 \vee \overline{x}_1\overline{x}_2\overline{x}_3\}, \{x_1x_2x_3 \vee \overline{x}_1\overline{x}_2x_3 \vee \overline{x}_1x_2\overline{x}_3 \vee x_1\overline{x}_2\overline{x}_3\},$$

$$\{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2x_3 \vee \bar{x}_1x_2\bar{x}_3\}, \{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2x_3\}, \{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2 \vee \bar{x}_2\bar{x}_3 \vee \bar{x}_1\bar{x}_3\},$$
$$\{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2 \vee \bar{x}_1\bar{x}_3\}, \{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2\}.$$

Таким образом, мы видим, что даже для $n=3$ (три бинарных фактора) количество типов взаимодействующих классов (8) заметно больше, чем для $n=2$ (2). Однако по виду функций, представляющих эти классы (хотя и записанных в экономном виде минимальных ДНФ), довольно трудно понять характер совместного действия факторов в них и каким-либо образом сравнить эти типы друг с другом.

Определение 2. Пусть $f(\alpha)=1$ для некоторого $\alpha \in \mathrm{B}^n$. Назовём *степенью взаимодействия факторов* $x_1,\ldots,x_n$ в отклике *f при значении уровней факторов* $x=\alpha$ число

$$\mu_f(\alpha)=\begin{cases}\min\left\{d(\alpha,\beta)\middle|\beta\in C_f,\ \beta\neq\alpha\right\}-1, & \text{если } \left|C_f\right|>1\\ n, & \text{если } \left|C_f\right|=1\end{cases}$$

Здесь $d(\alpha,\beta)$ – расстояние Хэмминга между точками $\alpha,\beta \in \mathrm{B}^n, C_f=\left\{\alpha\middle|f(\alpha)=1\right\}$. Из определения следует, что выполняется неравенство $0 \le \mu_f(\alpha) \le n$.

Теорема 2. В отклике *f* есть взаимодействие факторов $x_1,\ldots,x_n$ при значениях их уровней $x=\alpha$ тогда и только тогда, когда $\mu_f(\alpha)\ge 1$.

Понятие степени взаимодействия позволяет без детального анализа структуры отклика определить наличие или отсутствие взаимодействия (в смысле, сформулированном в Определении 1) простой проверкой неравенства $\mu_f(\alpha)\ge 1$. Отсюда также следует, что функции *f* с условием $\mu_f(\alpha)=0$ и только они представляют отклики *без взаимодействия* для *n* факторов при значениях их уровней $x=\alpha$.

Определение 3. Назовём *степенью взаимодействия факторов* $x_1,\ldots,x_n$ в отклике *f* число $\mu_f=\max\left\{\mu_f(\alpha)\middle|\alpha\in C_f\right\}$.

Очевидно, что степень взаимодействия μ_f инвариантна относительно действия группы симметрий булева куба, т.е. эта величина корректно определена на *классе* эквивалентных откликов.

Имеет место следующий аналог Теоремы 2 для степени взаимодействия.

Теорема 3. В отклике *f* есть взаимодействие факторов $x_1,\ldots,x_n$ тогда и только тогда, когда выполняется неравенство $\mu_f\ge 1$.

Таким образом, класс $\{f\}$ будет представлять взаимодействие факторов $x_1,\ldots,x_n$ тогда и только тогда, когда выполняется неравенство $\mu_f\ge 1$. Приведем значения степени взаимодействия для перечисленных в Теореме 1 откликов.

Теорема 4. При $n=2$ степень взаимодействия для типов с взаимодействием *n* факторов равна: 2 для $\{x_1x_2\}$ и 1 для $\{x_1x_2 \vee \bar{x}_1\bar{x}_2\}$. Для $n=3$ имеем:

Таблица 1.

Классы откликов	Степень взаимодействия
$\{x_1x_2x_3\}$	3
$\{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2\bar{x}_3\}$	2
$\{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2x_3 \vee \bar{x}_1x_2\bar{x}_3 \vee x_1\bar{x}_2\bar{x}_3\}$, $\{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2x_3 \vee \bar{x}_1x_2\bar{x}_3\}$, $\{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2x_3\}$, $\{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2 \vee \bar{x}_2\bar{x}_3 \vee \bar{x}_1\bar{x}_3\}$, $\{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2 \vee \bar{x}_1\bar{x}_3\}$, $\{x_1x_2x_3 \vee \bar{x}_1\bar{x}_2\}$.	1

Так как полную конъюнкцию $x_1x_2\cdots x_n$ можно считать откликом с самым сильным взаимодействием (аналог произведения предикторов в регрессионном анализе), а отклики без взаимодействия характеризуются условием $\mu_f = 0$, то из Теоремы 4 следует, что величина μ_f действительно является некоторой характеристикой силы совместного действия.

Обсуждение. Предложенная выше методика классификации откликов по степени взаимодействия построена на теории булевых функций и хорошо известных в ней понятиях, которые адаптированы к рассматриваемой задаче. Это позволяет не только ответить на поставленные выше вопросы, но и рассматривать эту теорию как специальную область приложений теории булевых функций с возможностью постановки новых задач, которые в чисто эпидемиологической постановке не могли бы появиться. Примером этого и является введение понятия степени взаимодействия.

ЛИТЕРАТУРА

1. Mackie J. L. Causes and conditions // Am. Philos. Q. 1965. V. 2. P. 245-255.
2. Lewis D. Causation // J. Philosophy. 1973. V. 70. P. 556-567.
3. MacMahon B., Pugh T.F. Causes and entities of disease // Preventive Medicine. Boston: Little Brown, 1967. P. 11-18.
4. Rothman K. Causes // Am. J. Epidemiology. 1976. V. 104(6). P. 587–592.
5. Miettinen O. S. Causal and preventive interdependence: Elementary principles // Scand. J. Work. Environ. Health. 1982. V. 8. P. 159-168.
6. VanderWeele T.J., Robins M. The identification of synergism in the sufficient-component-cause framework // Epidemiology. 2007. V. 18(3). P. 329-339.
7. VanderWeele T.J., Richardson T.S. General theory for interactions in sufficient cause models with dichotomous exposures // Ann. Statistics. 2012. V. 40. P. 2128-2161.
8. Панов В. Г., Нагребецкая Ю. В. Алгебраическая трактовка двухфакторной теории достаточных причин // Труды СПИИРАН. 2013. Т. 3(26). С. 277–296.
9. Панов В. Г., Нагребецкая Ю. В. Алгебраическая классификация совместного действия n бинарных факторов // Материалы IX Междунар. конф. «Системный анализ в медицине». Благовещенск. 2015. С. 31-34.
10. Panov V.G. and Nagrebetskaya J.V. Boolean algebras and classification of interactions in sufficient-component cause model // Int. J. Pure Appl. Math. 2015. V. 98(2). P. 239–259.
11. Panov V.G. and Nagrebetskaya J.V. Classification of combined action of binary factors and Coxeter groups // J. Discr. Math. Sci. & Cryptography. 2018. V. 21(3). P. 661–677.
12. Яблонский С. В. Введение в дискретную математику. М.: Наука, 1986.

E-mail: I.V.Nagrebetckaia@urfu.ru, vpanov@ecko.uran.ru