

УДК 519.8

СПОСОБ ВЫДЕЛЕНИЯ ГЛАВНЫХ РЕЗУЛЬТИРУЮЩИХ СОСТАВЛЯЮЩИХ
В АЛГЕБРАИЧЕСКОЙ МОДЕЛИ КОНСТРУКТИВНОЙ ЛОГИКИ

О.В. ХРОМУШИН

Тулское региональное отделение Академии медико-технических наук,
e-mail: oleg@khromushin.com

Аннотация: Предлагается способ выделения главных результирующих составляющих в алгебраической модели конструктивной логики, основанной на вычислении границы раздела по наибольшей разности между графиком распределения мощностей и прямой линией, соединяющей точки с максимальной и минимальной мощностями.

Ключевые слова: анализ, алгебраическая модель, способ.

THE METHOD OF SEPARATION OF MAIN RESULTANT COMPONENTS IN THE ALGEBRAIC
MODEL OF CONSTRUCTIVE LOGIC

O.V. KHROMUSHIN

Tula Regional Branch of Academy of Medical and Technical Sciences,
e-mail: oleg@khromushin.com

Abstract: This study is to provide a method of separation of main resultant components in the algebraic model of constructive logic that is based on the calculation of the interface according to peak-to-peak deviation between the power distribution graph and the straight line connecting the points of maximum and minimum power.

Key words: analysis, algebraic model, method.

В последние 15 лет *алгебраическая модель конструктивной логики* (АМКЛ) широко применяется для системного анализа в биологии, химии и здравоохранении [1-5]. Одним из наиболее сложных для аналитика вопросом применения АМКЛ является интерпретация модели, в том числе выделение главных результирующих составляющих.

Результирующие составляющие (импликанты), представляют собой переменные (факторы) с указанием пределов значений (объединенные через знак конъюнкции с другими факторами, в случае их совместного воздействия) и с указанной результирующей мощностью (W), по которой можно судить о степени их влияния на результат.

Мощность результирующей составляющей представляет собой число строк в исходной базе данных, которые удовлетворяют условию результирующей составляющей. Пределы определения каждого фактора алгоритм АМКЛ выбирает исходя из требования компактного представления результата.

Оценивая результат, аналитик, прежде всего, стремится выделить наиболее мощные (главные) составляющие. Границу, отделяющую главные результирующие составляющие от второстепенных можно провести следующими способами:

- вручную, на основании опыта аналитика и обоснованных доводов;
- по методике, основанной на определении границы раздела путем суммирования с накоплением мощности результирующих составляющих сверху вниз и снизу вверх с нахождением совпадающих точек в этих двух рядах или по наименьшей разности значений [6];
- по предлагаемой методике.

Основопологающая идея предлагаемого метода заключается в нахождении максимальной абсолютной величины разности, показанной на рис. 1 ($81 - 40 = 41$), при рассмотрении сверху вниз, как границы, отделяющей главные составляющие (слева на рис. 1) от прочих составляющих (справа на рис. 1).

Таким образом, необходимо вычислить минимальное значение разности:

$$W_i - \left[W_1 - \frac{W_1 - W_n}{n-1} (i-1) \right] = W_i - \frac{(n-i)W_1 + (i-1)W_n}{n-1}. \quad (1)$$

Обозначения:

1. W_i – мощность результирующей составляющей, где $W_1 = W_{\max}$, $W_n = W_{\min}$ (показано на графике синим цветом).
2. n – число результирующих составляющих (на графике $n=7$).
3. S_i – мощность результирующих составляющих, которые могли бы быть при линейном характере их убывания, где $S_1 = W_1$, $S_n = W_n$ (показано на графике коричневым цветом).
4. i – номер результирующей составляющей ($i = 1, 2, 3 \dots n$).

Практика выполнения аналитических расчетов с помощью АМКЛ, что в большинстве случаев предлагаемый способ можно применять для выделения главных результирующих составляющих.



Рис. 1. Иллюстрация предлагаемого способа выделения главных результирующих составляющих

Особенности предлагаемого метода:

Ситуация 1: График мощностей результирующих составляющих имеет более одной точки перегиба. В результате могут быть положительные значения разности (1), которые не могут исказить результат, поскольку предусмотрено нахождение максимальной разности. Если положительная область на графике будет преобладать, то надо вводить ограничения на использование предлагаемого метода.

Ситуация 2: График мощностей результирующих составляющих представлен линейной функцией. Теоретически результат может быть представлен составляющими с единичной мощностью, что обычно свидетельствует об отрицательном результате вычислений. В этих случаях данный способ не применим.

Ситуация 3: График мощностей результирующих составляющих имеет одну точку перегиба и соответствует рис. 1. Практика выполнения аналитических расчетов показывает, что такая ситуация чаще всего соответствует удачно выполненным аналитическим расчетам с большим числом случаев, соответствующих достижению цели, и достаточно большой базой сравнения (случаев, где цель не достигается).

Таким образом, ограничениями указанного способа являются наличие положительной области разностей (1) более установленного порога (предлагается 20%) или линейный характер графика мощностей (включая график с единичной мощностью).

Оценим точность предлагаемого способа на примере анализа смертности пешеходов с целью выявления угрозы жизни пешеходов от *дорожно-транспортных происшествий* (ДТП), характерных для Тульской области.

Характеристика массива исходных данных:

Источник информации – регистр смертности по Тульской области за 5 лет.

Число случаев – 2995 случаев смерти от внешних причин.

Случаи, характеризующие цель анализа – 1378 случаев смерти пешеходов.

Анализируемые факторы:

X1 – Цель, равная 1 для пешеходов (внешние причины).

X2 – Номер группы кодов травм (МКБ-X):

- 1 – "S00" - до "S10";
- 2 – "S10" - до "S20";
- 3 – "S20" - до "S30";
- 4 – "S30" - до "S40";
- 5 – "S40" - до "S50";
- 6 – "S50" - до "S60";
- 7 – "S60" - до "S70";
- 8 – "S70" - до "S80";
- 9 – "S80" - до "S90";
- 10 – "S90" - "S99.9";
- 11 – "T00" - до "T08";
- 12 – "T08" - до "T15";
- 13 – "T15" - до "T20";

14 – "Т20" - "Т98.3".

X3 – Возраст в годах (-1 – не известно, 0 – до года, 1 – от 1 до 2 лет, и т.д.).

X4 – Смерть произошла:

1 – от заболевания;

2 – от несчастного случая, не связанного с производством;

3 – от несчастного случая, связанного с производством;

4 – от убийства;

5 – от самоубийства;

6 – род смерти не установлен.

X5 – Смерть последовала:

1 – в стационаре;

2 – дома;

3 – в другом месте.

X6 – Признак город/село (1 – город, 2 – село).

X7 – Пол (1 – мужской, 2 – женский).

X8 – Образование:

1 – высшее;

2 – неоконченное высшее;

3 – среднее специальное;

4 – среднее общее;

5 – неполное среднее;

6 – начальное и ниже;

7 – неизвестно.

X9 – Семейное положение:

1 – состоял(а) в браке;

2 – никогда не состоял(а) в браке;

3 – вдов(а);

4 – разведен(а);

5 – неизвестно.

X10 – Месяц смерти.

Характеристика полученной математической модели:

– число результирующих составляющих – 789 шт. ;

– мощность первой результирующей составляющей – 31;

– мощность последней результирующей составляющей – 1;

– ступенчатый характер кривой распределения мощности (рис. 2);

– значительный (редко встречающийся в практике аналитических расчетов) участок кривой распределения мощности с малыми значениями.

Оценку точности предлагаемого способа производим в следующей последовательности:

1. Строим график разностей по формуле (1), показанный на рис. 3.

2. Находим на графике наименьшую величину (на рис. 3 показана точка 122 результирующей составляющей).

3. С помощью специальной аналитической программы AnAMCL строим графики влияния одного из факторов на результат с 122 и 789 результирующими составляющими [7]. Графики строятся путем изменения одного из факторов (в данном случае X10 – месяц смерти) при заданных значениях остальных факторов. При этом в качестве результата принимается суммарная мощность тех результирующих составляющих, в пределы определения которых попадают значения факторов.

4. С помощью программы AnAMCL находим максимально возможную суммарную мощность результирующих составляющих для двух вариантов: с 122 и 789 результирующими составляющими. При этом указанная программа пропускает через полученную математическую модель (как через фильтр) значения факторов каждого случая из базы, отыскивая вариант с максимальным значением суммарной мощности.

5. Осуществляем сравнение графиков (рис. 4 и 5) и значений суммарных мощностей.

Принятые условия для построения графиков (рис.4 и 5):

X2 =2 (номер группы кодов травм);

X3 = 80 лет;

X4 = 2 (смерть произошла от несчастного случая, не связанного с производством);

X5 = 3 (смерть последовала в другом месте);

X6 = 1 (город);

X7 = 1 (мужской);

X8 = 1 (образование высшее);

X9 = 1 (в браке);

X10 изменяется от 1 до 12 месяцев.

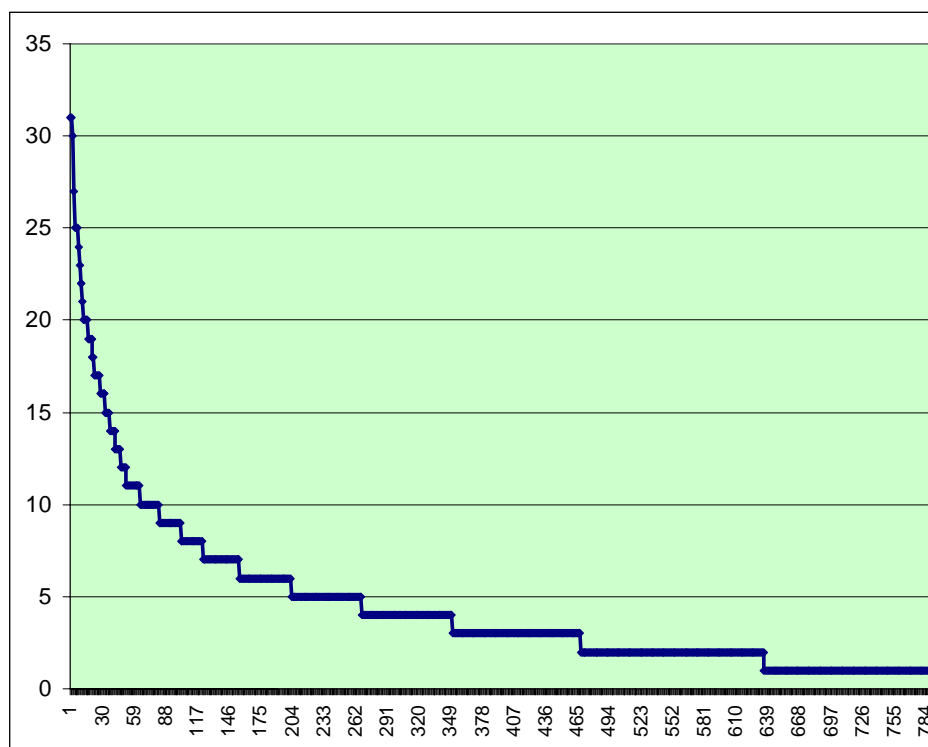


Рис. 2. Распределение мощностей результирующих составляющих

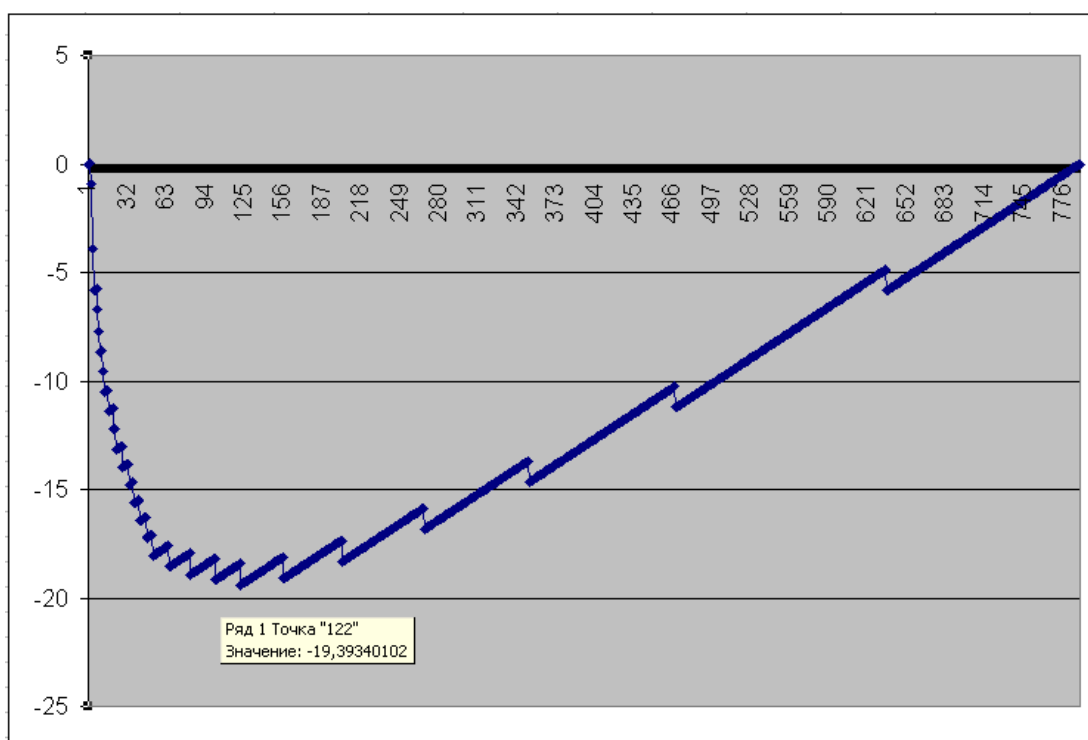


Рис. 3. Разности, подсчитанные по предлагаемой методике

Из полученных графиков (рис. 4 и 5) следует:

1. Суммарная мощность не изменилась.
2. Характер изменения существенно не изменился. Незначительные изменения коснулись начальной части графика (январь месяц) и величин изменений сентябрь – октябрь, март – апрель.
3. На рис. 5 просматривается постоянная составляющая $W=8$, которой нет на рис. 4.

ЗАВИСИМОСТЬ СУММАРНОЙ МОЩНОСТИ РЕЗУЛЬТИРУЮЩИХ
ИМПЛИКАНТ В ДИАПАЗОНЕ ИЗМЕНЕНИЯ ВЫБРАННОГО ФАКТОРА ОТ
МИНИМАЛЬНОГО ДО МАКСИМАЛЬНОГО ЗНАЧЕНИЯ

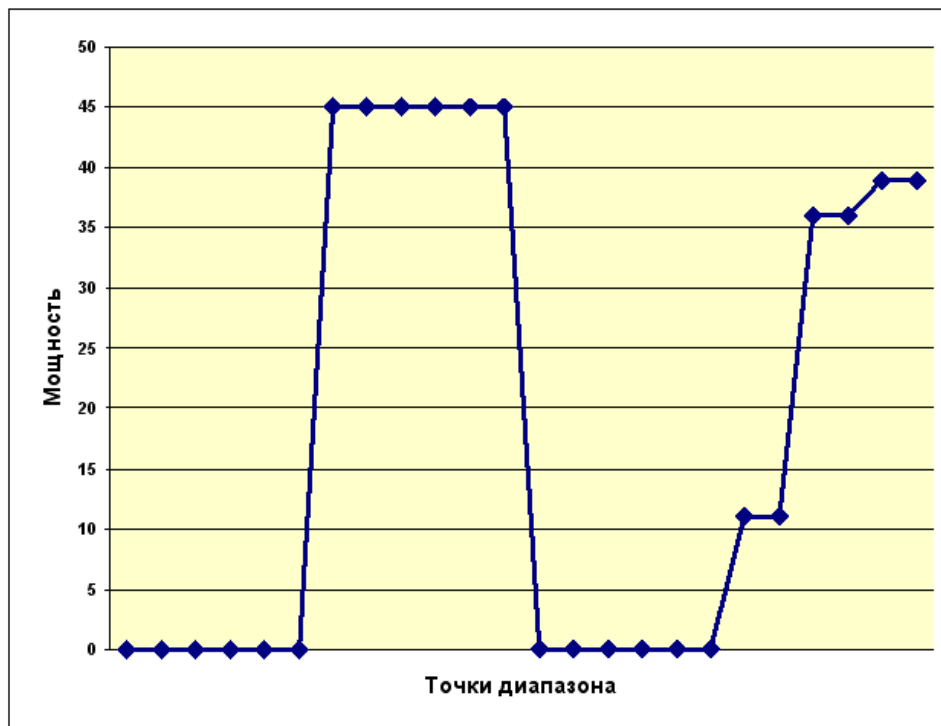


Рис. 4. График для X10 (месяц смерти) для модели с 122 результирующими составляющими, $W_{\max}=210$

ЗАВИСИМОСТЬ СУММАРНОЙ МОЩНОСТИ РЕЗУЛЬТИРУЮЩИХ
ИМПЛИКАНТ В ДИАПАЗОНЕ ИЗМЕНЕНИЯ ВЫБРАННОГО ФАКТОРА ОТ
МИНИМАЛЬНОГО ДО МАКСИМАЛЬНОГО ЗНАЧЕНИЯ

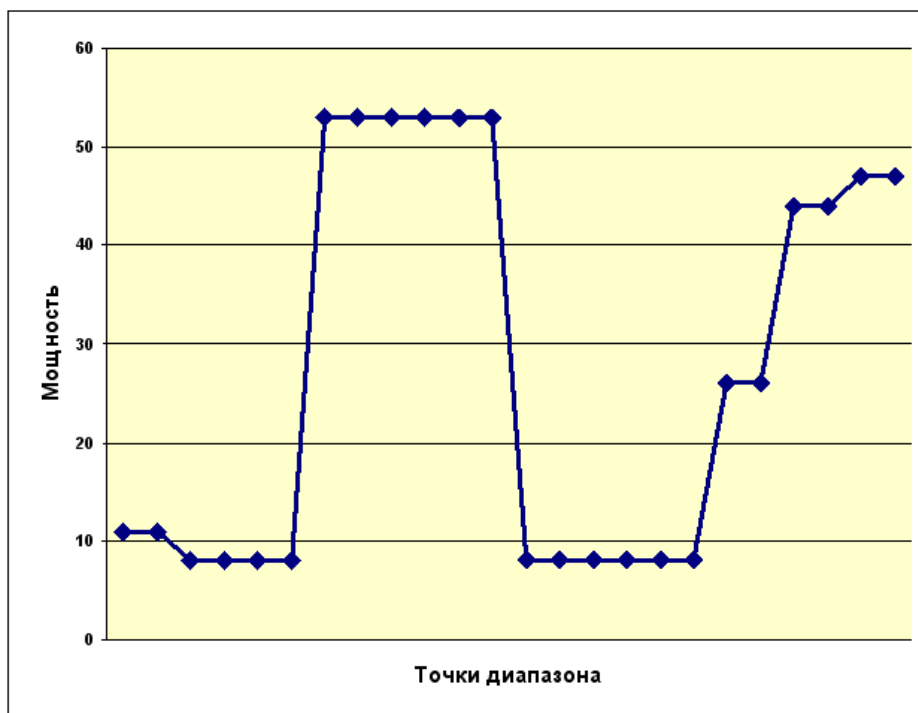


Рис. 5. График для X10 (месяц смерти) для модели со всеми результирующими составляющими (789 шт.),
 $W_{\max}=210$

Оценивая точность Θ , необходимо учитывать суммарную мощность. Поскольку суммарная мощность не изменилась, можно оценку производить постоянной составляющей: $\Theta = 8/210 = 0,038$ (3,8%).

Полученная оценка точности является приемлемой для большинства аналитических расчетов.

Выводы. Предлагаемый способ целесообразно использовать в практике интерпретации результата алгебраической модели конструктивной логики, что позволит расширить возможности аналитика.

Литература

1. *Хромушин, В.А.* Обзор аналитических работ с использованием алгебраической модели конструктивной логики / В.А. Хромушин, А.А. Хадарцев, О.В. Хромушин, Т.В. Честнова. – Тула: Вестник новых медицинских технологий. Электронное издание, (2011. – N1, публикация 3–2), <http://www.medtsu.tula.ru/VNMT/Bulletin/E2011-1/LitObz.pdf>.

2. *Честнова, Т.В.* Контекстно-развивающаяся база данных для логической интеллектуальной системы, используемой в здравоохранении / Т.В. Честнова, В.Н. Щеглов, В.А. Хромушин // Эпидемиология и инфекционные болезни. – 2001. – N4. – С. 38–40.

3. *Хромушин, В.А.* Нечеткая алгебраическая модель конструктивной логики / В.А. Хромушин, Е.И. Минаков, В.А. Бархоткин, О.В. Хромушин, В.Ф. Бучель. – Тула: Вестник новых медицинских технологий. – 2012. – N1. – С. 36–38.

4. *Хромушин, В.А.* Упрощенный вариант алгебраической модели конструктивной логики / В.А. Хромушин, Е.И. Минаков, О.В. Хромушин, В.А. Бархоткин, С.Н. Гонтарев. – Тула: Вестник новых медицинских технологий. – 2012. – N1. – С. 44–47.

5. *Хромушин, В.А.* Алгоритмы и анализ медицинских данных / В.А. Хромушин, А.А. Хадарцев, В.Ф. Бучель, О.В. Хромушин. Учебное пособие. – Тула: Изд-во «Тульский полиграфист», 2010. – 123 с.

6. *Хромушин, В.А.* Обобщенная оценка результирующей алгебраической модели конструктивной логики / В.А. Хромушин, В.В. Махалкина // Вестник новых медицинских технологий. – Тула: ТулГУ. – 2009. – N 3. – С. 39–40.

7. *Хромушин, В.А.* Алгоритм и программа анализа результирующих импликант алгебраической модели конструктивной логики / В.А. Хромушин, О.В. Хромушин, Е.И. Минаков // XXXXVI научно-практическая конференция профессорско-преподавательского состава ТулГУ "ОБЩЕСТВЕННОЕ ЗДОРОВЬЕ И ЗДРАВООХРАНЕНИЕ: ПРОФИЛАКТИЧЕСКАЯ И КЛИНИЧЕСКАЯ МЕДИЦИНА": Сборник статей. – Тула, 2010. – С. 138–148.