

# Mathematics for logical computer-aided analysis of decisions in agromarketing

Boris Kulik<sup>1</sup>, and Alexander Fridman<sup>2,\*</sup>

<sup>1</sup>Institute of Problems in Mechanical Engineering of RAS, 199178 St. Petersburg, Russia

<sup>2</sup>Institute for Informatics and Mathematical Modelling, Kola Science Centre of RAS, 184209 Apatity, Murmansk reg., Russia

**Abstract.** Collecting data on possible solutions and comparing them should be an integral part of any research in agromarketing. Publications on this topic offer a number of quantitative methods for carrying out such an analysis, but their reliability depends significantly on the accuracy of the initial data, which is usually low. This article describes a logical framework for detecting uncertainties, inconsistencies, and contradictions in the information regarding possible decisions with using our previously developed  $n$ -tuple algebra. In our opinion, this allows to computerize an important part of planning tasks in agromarketing and provides good means to objectify and enrich results of any analysis of decision options.

## 1 Introduction

Many tasks related to development of the agro-industrial complex (AIC) require marketing study to select a preferred option for their implementation. In the field of strategic planning, such tasks include zoning of agricultural crops [1], selection of cattle for beef, dairy and breeding stock [2], etc. Importance of using artificial intelligence (AI) methods in marketing research is discussed by many experts (for example, [3]). The following goals of applying cognitive marketing methods are proposed:

- identify unmet customer needs and upgrade products and services;
- attract more customers and increase profits with customized products, services, benefits, offers and pricing;
- get the most out of cooperation with existing customers and partners by optimizing investments into strengthening relationships.

In our opinion, the listed and many other tasks of marketing in the AIC admit formal analysis, which should include studying of inevitable uncertainties, inconsistencies and contradictions in information about the food market. Formalization must be intelligent [4].

The mathematical logic used in AI is a universal and widespread tool for performing this type of analysis, but it is poorly suited for analyzing uncertainties and inconsistencies. That is why we further propose a logical apparatus for discovering such informational mismatches by using our earlier developed  $n$ -tuple algebra (NTA) (e.g., [5-7]).

---

\* Corresponding author: [fridman@iimm.ru](mailto:fridman@iimm.ru)

To model and analyze uncertainties and incorrectness in AI, non-classical, in particular, non-monotonic, modal and paraconsistent logics are often used [8, 9]. NTA includes techniques to discover and study such cases within the frames of classical logic.

NTA is based on a new way of representing Cartesian products (CPs) of sets. Instead of the traditional notation of CPs as, for instance,  $A \times B \times C$ , a new structure is used, namely the **C-n-tuple**  $P[XYZ] = [A B C]$ , where  $P$  is the name of the relation,  $[XYZ]$  is the **relation diagram**,  $X, Y, Z$  are **attributes**,  $A, B, C$  are **components** (subsets of the corresponding attributes). Since CP is not a universal way of representing all relations, new types of models called **NTA-objects** were proposed for relations. In addition to C-n-tuples, they are C-systems (unions of C-n-tuples with the same relation diagram), **D-n-tuples** (complements of C-n-tuples) and **D-systems** (complements of C-systems).

C-systems are represented in the form of matrices bounded by square brackets (for example,  $Q[XYZ] = \begin{bmatrix} A & B & C \\ D & E & F \end{bmatrix}$ ); D-n-tuples look like tuples enclosed in inverted square

brackets (for instance,  $\bar{P}[XYZ] = ] \bar{A} \bar{B} \bar{C} [$  is the complement of the C-n-tuple  $P[XYZ]$ , where  $\bar{A} = X \setminus A$ ,  $\bar{B} = Y \setminus B$ , and  $\bar{C} = Z \setminus C$ ); D-systems are written as matrices put into

inverted square brackets (thus,  $\bar{Q}[XYZ] = \begin{bmatrix} X \setminus A & Y \setminus B & Z \setminus C \\ X \setminus D & Y \setminus E & Z \setminus F \end{bmatrix}$  is the complement of the

C-system  $Q[XYZ]$ ). In NTA, algorithms have been developed for performing all operations and comparisons of algebra of sets with any NTA-objects, as well as ways of representing problems and methods of mathematical logic with structures and algorithms of NTA.

## 2 Materials and Methods

### 2.1 Variants of values

Let us consider how NTA helps to express uncertainties in the form of variants of values. In predicate calculus, only two ways of representing variables in formulas are allowed, namely *complete ambiguity*, when a variable name is used, and *complete unambiguity*, when a variable is replaced by a constant or a function of a constant. To represent intermediate states, you need to complicate the formula. At the same time, NTA has the ability to set options for values directly in the structures themselves. Let us look at an example. Suppose, three experimental sorts of wheat (sorts A, B and C) were planted on three plots, but the records of what was planted where were lost. Only some information about this has survived, in particular, the following "There is not Sort B on the first plot, there is not Sort A on the second plot". How can this restriction be expressed in NTA language?

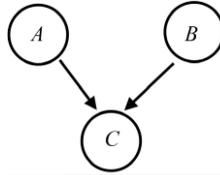
A single C-n-tuple is sufficient for this:  $Q[P_1P_2P_3] = [\{A, C\} \{B, C\} *]$ .

Here  $P_i$  are attributes with possible list of values (**domain**)  $\{A, B, C\}$ ; the first and the second components are given by two possible variants of values, for the third (**dummy**) component any value from the domain is allowed.

### 2.2 Probabilistic analysis

In AI, Bayesian networks [8] are used as the main method of probabilistic analysis, with the help of which probabilistic estimates of cause-and-effect relations are calculated. A Bayesian network is a directed acyclic graph whose vertices represent events with a given probability, and the links (arcs) between the vertices are described using tables of conditional probabilities.

Consider an example of a network (see Figure 1), in which the probability of the vertex  $C$  being in different states ( $c_k$ ) depends on the states ( $a_i, b_j$ ) of the “parent” vertices  $A$  and  $B$ . The probability of the state  $c_k$  of the vertex  $C$  is calculated using the expression  $p(c_k) = \sum_i \sum_j p(c_k | a_i, b_j) \times p(a_i, b_j)$  where  $p(c_k | a_i, b_j)$  is the conditional probability of the state  $c_k$  depending on the states  $a_i$  and  $b_j$ .



**Fig. 1.** Bayesian network example

If we assume that each event  $x$  in the given system can take only two states (say, *true* ( $x_1$ ) and *false* ( $x_0$ )), then even for this simple case, a complete description of the Bayesian network requires specifying six values: two state probabilities *true* ( $a_1$  and  $b_1$ ) for events  $A$  and  $B$  and four conditional probabilities for the event  $c_1$  with the following combinations of values of states for “parent” events:  $(a_1, b_1)$ ,  $(a_0, b_1)$ ,  $(a_1, b_0)$  and  $(a_0, b_0)$ .

Consider some possibilities of probabilistic modelling of dependent events with using NTA. In [6], a method of probabilistic modelling is described for an arbitrary logical formula, which allows to express the relations between the states of the variables of this formula in the form of a probabilistic function. In particular, for the example under consideration, the following logical formula can be proposed:  $F = (A \wedge B) \supset C$ . Provided that each variable has two states, the following probability function can be derived [6]:

$$p(F) = 1 - p(a_1) p(b_1) + p(a_1) p(b_1) p(c_1).$$

The obtained formula shows that the complete NTA description of this system requires to set only three values of the state probabilities.

If, with the same scheme in Figure 1, the semantics of the problem under consideration does not correspond to the logical formula  $F = (A \wedge B) \supset C$ , another formula can be applied, for instance,  $F = (A \vee B) \supset C$ . It is transformed into the disjunctive normal form (DNF)  $F = (\neg A \vee C) \wedge (\neg B \vee C)$ . Then NTA methods described in [6] yield the following probability function:

$$p(F) = 1 - p(c_0) + p(a_0)p(b_0) p(c_0).$$

Here, as in the previous example, the complete description of the system needs setting of only three values of the state probabilities. In addition to the considered dependencies, other logical formulas can be used for this system, which do not necessarily correspond to the network in Figure 1.

### 2.3 Clustering of objects

Let us consider a class of objects whose values of all attributes can be ordered, and they contain not only quantitative attributes (time, distance, price, etc.), but also qualitative ones, which are given by names (sorts of cereals and so on). For such attributes, we can choose (for example, based on expert judgments) a measure of preference. Studies have shown that such structures (*clusterizable NTA objects* [7]) have new properties.

These objects can also have attributes whose values are specified as a finite set of intervals. For this, the *interval quantization method* proposed in [5, 7] is used, with the help

of which the system of arbitrary intervals on the coordinate axis is expressed in the form of a linearly ordered set of quanta. For example, the intervals on the axis of real numbers (2.1 - 7.6), (3.5 - 6.5), (4.8 - 9) can be divided into quanta by writing the boundaries of the intervals in ascending order: 2.1, 3.5, 4.8, 6.5, 7.6, 9. For this sequence, elementary intervals (quanta) will be: (2.1 - 3.5), (3.5 - 4.8), (4.8 - 6.5), (6.5 - 7.6), (7.6 - 9) with names (numbers) 1, 2, 3, 4, 5. Then each of the original intervals can be represented as a set of quanta. In particular, the interval (2.1 - 7.6) is formed by the set of quanta {1, 2, 3, 4}, and the interval (4.8 - 9) consist of the set {3, 4, 5}.

The proposed transformations of attributes make it possible to divide the objects under study into clusters, to calculate the degrees of connectivity and the distance between them [7]. The resulting dependencies expand the scope of solvable classification problems.

## 2.4 Analysis of contradictions

Studies of the food market, taking into account the many shortcomings in management and competition between producers, show the need for a logical analysis of this phenomenon in order to identify possible contradictions.

In mathematical logic, a contradiction is recognized as a situation in which a logical formula is identically false [10]. Many works on the foundations of mathematics and mathematical logic are devoted to the analysis of contradictions.

In our works [5, 11], we have proposed several methods for elimination of contradictions. In particular, one of them was developed in the study of properties of presupposition [11]. A *presupposition* is a statement that is implied (or perceived as true) when the main statement is actualized, while the denial (or falsity) of the main statement does not violate the truth of the presupposition. Another feature is that falsity or inconsistency of the presupposition entails the loss of the meaning of the main statement.

For example, the sentence *Elimination of the consequences of a technogenic catastrophe deserves a high assessment* implies that *Elimination of the consequences of a technogenic catastrophe was carried out* – this is a presupposition. It is clear that the denial of the main statement (*Elimination of the consequences of a technogenic catastrophe was recognized as unsatisfactory*) does not affect the truth of the presupposition. If we assume that the presupposition is false, then assessment of actions does not make any sense.

In works on logic and linguistics [12, 13], the presupposition is often expressed as a logical formula  $(S \supset P) \wedge (\neg S \supset P)$  where  $S$  is the main statement (*assertion*), and  $P$  is a presupposition. However, in many cases, a presupposition means a prerequisite: often events expressed in assertion occur after events expressed in presupposition. However, if you depict this by using the formula  $(P \supset S) \wedge (P \supset \neg S)$ , you obtain the equivalent formula  $\neg P$ , which means the unconditional falsity of the presupposition. Conversely, an informal analysis of many specific presuppositions shows that there really is no paradox. A typical example provides a clue to solving this problem.

Consider the phrase ( $S$ ) *This product is not in great demand*. Presupposition for it is the sentence ( $P$ ) *This product has been put up for sale*. In order to obtain an estimate of demand, the product must be put up for sale, i.e.  $P$  is quite valid here as a prerequisite. However, this does not take into account possible reasons for low or, conversely, high demand (advertising, product quality, etc.). These reasons can only be displayed by using a new logical variable we called “*assertion switch*”. Let us formulate the following hypothesis.

*Hypothesis 1.* If  $P$  is a presupposition of the sentence  $S$ , then an *assertion switch*  $R$  exists and can be found by the following properties: the expression  $P \wedge R$  is the prerequisite of the sentence  $S$ , and the expression  $P \wedge \neg R$  is the prerequisite of the sentence  $\neg S$ .

Taking this into account, the following consistent logical model of the presupposition is proposed:  $((P \wedge R) \supset S) \wedge ((P \wedge \neg R) \supset \neg S)$ . Analysis shows that there are no dummy variables in this formula, i.e., all variables in the main sentence ( $S$ ), presupposition ( $P$ ), and switch ( $R$ ) are significant. In addition, this formula does not initiate any paradox, since the truth of the presupposition does not entail the falsehood of the formula.

In knowledge bases, an analogue of the obtained logical model of the presupposition is an anomaly of contradiction [14]. To correct this anomaly, it is usually recommended to eliminate one of the conflicting rules. However, the approach proposed in the analysis of the presupposition allowed us to develop an algorithm for elimination of the contradiction anomaly by searching for additional parameters similar to the assertion switch [11].

### 3 Results and Discussion

This report proposes NTA-based mathematical methods to analyze the following uncertainties and incorrectness in knowledge: uncertainties in the form of variants of values; probabilistic modeling; reasoning with paradoxes and contradictions; ordering and clustering of objects with multiple and heterogeneous attributes.

Main results of our research, which can be useful in marketing, are as follows.

If a marketer needs to analyse alternative options, i.e., uncertainties are expressed in the form of variants of values, NTA structures can model these options directly (Section 2.1).

If it is better to apply probabilistic models, NTA-based probabilistic analysis allows to significantly reduce the amount of data required for a complete description of the system in comparison with probabilistic modelling by means of Bayesian networks. In addition, the use of NTA in probabilistic modelling provides expanding the options for structures and models compared to Bayesian networks (Section 2.2).

If it is required to divide marketing objects into classes, within which the similarity of objects is stronger than in different classes, clustering methods are applicable (Section 2.3).

Last but not least, sometimes it is possible to reveal and avoid contradictions by finding an assertion switch that clarify sources of the problem under study (Section 2.4).

Due to the limited size of this paper, it is not possible to give examples of solving agromarketing problems, therefore, we will simply list some (far from all) tasks of the agro-industrial complex, for which the described methods look promising:

- zoning of agricultural crops [1, 15];
- selection of cattle for meat, dairy and breeding livestock [2];
- environmental monitoring for the purpose of tracking and analyzing the state of rural areas in terms of the level of technogenic pollution [16];
- development of a personalized virtual advisor for buyers [17]
- synthesis of operational and strategic decisions throughout the entire cycle of the content plan [18].

### 4 Conclusion

The results of this study show that the classical approach to modeling some uncertainties of knowledge systems makes it possible to integrate various methods for solving such problems by means of our earlier developed  $n$ -tuple algebra.

In this paper, we propose an intelligent logical apparatus for discovering uncertainties, inconsistencies and contradictions in marketing information as well as for clustering objects with numerous and various attributes. In our opinion, this technique allows to computerize an important part of planning in agromarketing and provides a good means for objectifying and enriching results of any analysis of decisions options.

This work was supported in part by the Russian Foundation for Basic Research (grants 18-07-00132, 18-01-00076, 18-29-03022, and 19-08-00079).

## References

1. Yu.A. Tsyarkin, A.N. Lyukshinov, N.D. Eriashvili. *Agromarketing and consulting*. M.: UNITI (2000) (in Russian)
2. G.A. Bushova. *Breeding of dairy cattle for a complex of economic and biological characteristics*, diss. according to VAK RF 06.02.07, Ph.D. (Biol.) (2013) (in Russian)
3. *The age of machines: how artificial intelligence is changing* <https://www.sostav.ru/publication/professionalny-41131.html> (in Russian)
4. J. Bouma et al., *Advan. in Agron.*, **113**, 293-323 (2011)
5. B. Kulik, A. Fridman. *N-ary Relations for Logical Analysis of Data and Knowledge*. IGI Global (2017)
6. B.A. Kulik, *J. Comp. and Syst. Sci. Intl.*, **46(1)**, 111-120 (2007)
7. B.A. Kulik, A.Ya. Fridman, *Methods of Clustering on Logical Models. Pattern Recogn. and Image Analysis*, **30(2)**, 201-208 (2020)
8. S. Russel, P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. (Prentice Hall, 2009)
9. A. Thayse, P. Gribomont, G. Hulin et al. *Approche logique de l'intelligence artificielle, vol.2. De la logique modale à la logique des bases de données* (Paris, Dunod, 1989)
10. E. Mendelson, *Introduction to Mathematical Logic*. 6th ed. (Taylor & Francis, 2015)
11. B. Kulik, A. Fridman, *CEUR Workshop Proceedings*, **2303**, 163-177 (2018)
12. P. Strawson, *Introduction to Logical Theory* (London, 1952)
13. B. van Fraassen, *J. of Philos.*, **65(5)**, 136-152 (1968)
14. F. Harmelen, *Decis. Supp. Syst.*, **21(4)**, 271-280 (1997)
15. P.T. Mburu, and S.K. Massimo, *J. of Appl. Sci.*, **5**, 1133-1141 (2005)
16. N. Dayalan, N. Triuchirappalli-Tamil, *Transactions*, **40(2)**, 271-284 (2018)
17. *Artificial intelligence in marketing: the beginning of a new era* <https://www.cossa.ru/trends/198226/> (in Russian)
18. *Artificial intelligence in digital marketing: how the work of a marketer is changing in 2020.* <https://vc.ru/marketing/107257-iskusstvenny-intellekt-v-digital-marketinge-kak-menyaetsya-rabota-marketologa-v-2020-godu> (in Russian)