Some Mathematical and Practical Aspects of Decision-Making Based on Similarity

Peeter Lorents[1(✉)] and Maryna Averkyna[1,2(✉)]

[1] Estonian Business School, A. Lauteri, 3, Tallinn, Estonia
peeter.lorents@ebs.ee, maryna.averkyna@oa.edu.ua
[2] The National University of Ostroh Academy, Seminarska, 2, Ostroh, Ukraine

Abstract. One type of decision-making processes, which is often applied, based on the similarity of situations and developments. This study examines some approaches to addressing the similarities of situations and developments, including structural similarity and descriptive similarity. Structural similarity and descriptive similarity have been linked in many ways. One of these ways is based on theorems proven in algebraic systems and universal algebra theories. The authors point out that in order to assess two sets of descriptive similarity, it is first necessary to make descriptions of both sets, which must consist of relevant statements. The application of descriptive similarity in the process of managing the development of public transport systems in small towns in Estonia and Ukraine is considered. The authors presented the algorithm of decision-making process' method. The approach how to apply the descriptive similarity between Estonian small towns and their public transport systems, and the small town of Ostroh from Ukraine is proposed. Some concrete examples and derived preliminary conclusions is presented.

Keywords: Situations and developments as systems · Types of systems similarity · Structural similarity · Descriptive similarity · Numerical evaluation of descriptive similarity · Comparison of plausibility based on numerical estimates of similarity · Evaluation of the descriptive similarity · Public transport systems in the cities surveyed

1 Introduction

The purpose of this work is to clarify certain ways to obtain numerical estimates of the similarities between situations and developments to implement these assessments to manage situations and developments. The reason for this is the fact that people often make decisions based on the degree of to their situation similarity (or development) compared to some other known situation (or development). This can be illustrated by the following scheme:

(I) The decision-maker examine the current situation (development);
(II) Decision-maker finds sufficiently similar and already known situations (developments);

C. Stephanidis (Ed.): HCII 2019, LNCS 11786, pp. 168–179, 2019.
https://doi.org/10.1007/978-3-030-30033-3_13

(III) The decision-maker examines what was decided in the case of these situations (developments): In one case, for example, D1, in another case D2;
(IV) The decision maker will know that the D1 implementation was more positive than the result for D2;
(V) The decision-maker decides to do what is more like D1 because he believes that in this case, the result is similar (hence positive) to what was achieved with D1 in a known situation (development).

Here, it seems that the more experienced and more successful decision-makers do not rush to recommend D1, which gave positive results if it is not convinced that the situation (development) in which D1 was accepted is sufficiently similar to what the decision-maker is doing in this case. Now we inevitably come to the questions: what is this similarity and is it possible - and if so, how to assess in absolute terms the degree of similarity? There are several ways to study the similarities between situations and developments.

For one approach, it is first necessary to consider situations and developments as so-called structured sets or systems (in algebraic terms, see for example Lorents, Matsak [4]. It is then possible to study the similarity of situations (developments) using homomorphism of algebraic systems. In algebraic system theory and its applications, the observation of systems similarity is usually limited to proving the homomorphism of these systems, in particular the presence of isomorphism (or the lack thereof). At the same time, it has not been discussed what the rate of homomorphism could be and how to calculate it. (Homomorphism or algebraic similarity of finite systems can be numerically estimable by a method developed by P. Lorents only last year – 2018. This research is on publication.)

In the second approach to this work, the description of situations (developments) is based on claims. In this case, descriptions of the situations (developments) are a *set of claims*. In this context, situations (developments) are assumed to be similar if there are similar sets of claims. The numerical estimation of the similarity of the sets is based on a method analogous to that described in the works published in 1901 by Swiss botanist and plant physiologist Paul Jaccard [2, 3]. Very significant and substantive difference here is that the calculation of the Jaccard coefficient is based on the *equal elements* of the two sets. Nevertheless, we rely on *equalized elements* to calculate the similarity coefficient. One relevant approach is presented by Lorents P., Matsak E., Kuuseok A., Harik D. in the work published in 2017 [5].

Of course, the set and set of statements describing this set are not the same. But often we have no other practically usable ways to study the sets, relying on their descriptions. Nevertheless, there is a certain relationship between the set of structures, and the number of statements used to describe them. Indeed – the structural similarity and descriptive similarity are linked in many ways. One of these ways is based on theorems proven in algebraic systems and universal algebra theories (see, for example, Cohn [1], Maltsev [7]), according to which:

– in the case of two isomorphic (or “perfectly similar”) systems, all claims that are represented by the corresponding formulas and are correct in one system - are also correct in another system, and all claims that are correct in another system are correct in the first system

– if one system is homomorphic (or "reasonably similar" to another system), all the claims that can be made with the so-called positive formulas and that "fit" in the first system - will also "fit" in another system (but not always the other way around - what is right in another system may not always be right in the first system).

It is useful to use during decision-making process in order to implement Decision Support Systems in management urban public transportation, and it is important to develop making-decision approach. Szűts I., András B. emphasized the main approaches in decision-making: the psychological approach, the classical economy, and the administrative models approach the reality. However, these approaches have limitation. 'The psychological approach application in practice does not serve reliable result. It is difficult to apply model real situation according to the classical economy approach. Administrative models allow decision makers to solve the actual issue(s), and they have not alternative solution' [8]. It is necessary to point out, that descriptive similarity can be as a basic in the decision-making process management urban public transportation.

In the present paper, we will first discuss the concept of descriptive similarity and the calculation of the appropriate numerical estimate. We will then study how to apply the descriptive similarity between Estonian small towns and their public transport systems, and the small town of Ostroh from Ukraine. We present some concrete examples and derived preliminary conclusions. The further aim is to explain how sensible it is to implement solutions in other cities.

1.1 Descriptive Similarity and Its Numerical Assessment

In order to assess two sets of descriptive similarity, it is first necessary to make descriptions of both sets, which must consist of relevant statements. Next, it is necessary to clarify what statements from one description and another description can be considered equivalent. There are now three sets of claims:

– A is a set whose elements are statements from the first description;
– B is a set whose elements are statements from another description;
– C is a set of elements that are ordered pairs, where the first position of the pair has the claim A, the second position has the claim B, and these claims have been equalized by each other.

Note. Sometimes it is quite useful to find a common, both-for-one formulation for the claims that are equalized to one another. It is not forbidden, however, that one of these two formulations be suitable.

We agree that we denote the number of elements of the final set H with the symbol E(H). In order to calculate the coefficient of descriptive similarity between the sets A and B, we use the formula, if it is known, which claims are in this case equated with each other (or set C):

$$\mathrm{Sim_C(A,\ B)} = \mathrm{E(C) : [E(A) + E(B) - E(C)]}. \tag{1}$$

Example. We consider two families F' and F''. Describe the families of the statements S' i and S'' p to these families.

Family F': Living in the countryside. They have their own house. Families have four children. Father is working. Mother is home. There is a big garden around the house. There are two big dogs in the garden. The family has two cars. One usually runs a father, mother with another mother.

Family F'': There are five children in the family. Living in the city. They have a large apartment. Mom goes halfway at work. Grandmother helps deal with children. Father works as a deputy director at a large company. The family has three cars. One uses a grandmother, the mother and the third father travel to the other (Table 1).

Table 1. Families' statements.

The first family describes the claims	Equivalent claims			The other family describes the claims
	Wording from the first	Wording from the second	The wording for both	
Living in the countryside				
	They have their own house	They have a large apartment	They have their own home	
	Families have four children	Peres on viis last	Families have five children	
	Father is working	Father … as deputy director	Father is working	
Mother is home				
There is a garden around the house				
In the garden … two big dogs				
	The family has two cars	The family has three cars	Perel has several cars	
	With one … father, mother with another	One uses grandma, … and third father	Every adult has a car	
				Mom goes to work …
				With children … grandma

We will calculate the similarity rating:

$$5 : [9 + 8 - 5] = 5 : 12 \approx 0.42$$

Perhaps in this case there is an assessment that could be characterized by words: not very high or very low. Or by the words: rather low, than high. Alternatively, just words: not so small.

1.2 Some Explanations of Assessing the Similarity of Sets

We mentioned above that there is an important and substantive difference between the similarity estimation used here and the Jaccard coefficient. Let us explain this in some detail.

Let E(H) be the number of elements of a set H. Let us have two sets A and B.

- Relying on the ***same*** elements let us calculate the Jaccard's coefficient [2, 3]:

$$\mathrm{Sim_J(A, B)} = \mathrm{E(A \cap B) : [E(A) + E(B) - E(A \cap B)]} \tag{2}$$

- Relying on the ***equated (matched)*** elements let us calculate the Lorents coefficient (Lorents 2017):

$$\mathrm{Sim_{LT}(A, B)} = \mathrm{E(equ_T(A, B)) : [E(A) + E(B) - E(equ_T(A, B))]} \tag{3}$$

where **T** this is the way of equalization $\mathbf{equ_T(A,B)}$ is such a set, where x $\in$ A or y $\in$ B are belonging to i*n case, if* x and y **are *equated***. NB! **We expect**, that

- not any two elements from the set A are equatable with each other
- not any two elements from the set B are equatable with each other
- each element from the set A can be equated with at most one element from the set B and vice versa.

In case we believe that A = B.

Important Note 1. Equating *can* – but does *not necessarily have to* be guided by identity.

Example. Let's observe two sets of shapes:

A = {Δ,□,×,¬,♦,°,+} and B = {+, ×,∇,◊,⌉,●,o,□}. Apparently, in the minds of many, these sets have only two common elements: □ and × . Nevertheless, there are enough people who agree to equate also the following elements

Δ and ∇, + and +, ¬ and ⌉, ° and o. This is a way of equating T (for those people).

Let's calculate similarity coefficients:

$$\mathrm{Sim_J(A, B) = E(A \cap B) : [E(A) + E(B) - E(A \cap B)] = 2 : [7 + 8 - 2] = \mathbf{2} : \mathbf{13}}$$
$$\mathrm{Sim_{LT}(A, B) = E(equ_T(A, B)) : [E(A) + E(B) - E(equ_T(A, B))] = 6 : [7 + 8 - 6] = 6 : 9 = \mathbf{2} : \mathbf{3}!}$$

Important Note 2. The recent example is one confirmation of the fact that the assessment of similarity depends on how the equalization has taken place! Therefore, it is very important to exhibit in full what elements are equalized to each other. To some extent, a similar phenomenon (depending on how the similarity is determined by the way of identification) can also be observed for structural similarity. For example (by Lorents), if we compare fairly simple systems M ‘and M”, where: M’ elements are 2, 3, 4, 5, 6 and the relationship between them is “… is non-trivial multiplier for …”; the elements of M” are shapes O, □, Δ, and the relationship between them “… is more angular than …” – then we can make sure that there are 150 matches between these systems, 18 of which are suitable for homomorphism, and there are several quite similar numerical estimates of similarity (e.g. 0.2 and 0.5!). However, this means that if we limit ourselves to identifying only whether the systems are homomorphic or not homomorphic, we may not notice many aspects. This situation is somewhat reminiscent of what we can notice or not notice in a person’s photo, depending on the angle at which it is made.

2 Methodic Approach to Decision-Making Process’ in the Public Transportation

It is necessary point out, that descriptive similarity is the basis for developing decision-making process’ method in the public transportation (Fig. 1). This method will show relevance of studying and implementing the experience of managing urban public transport.

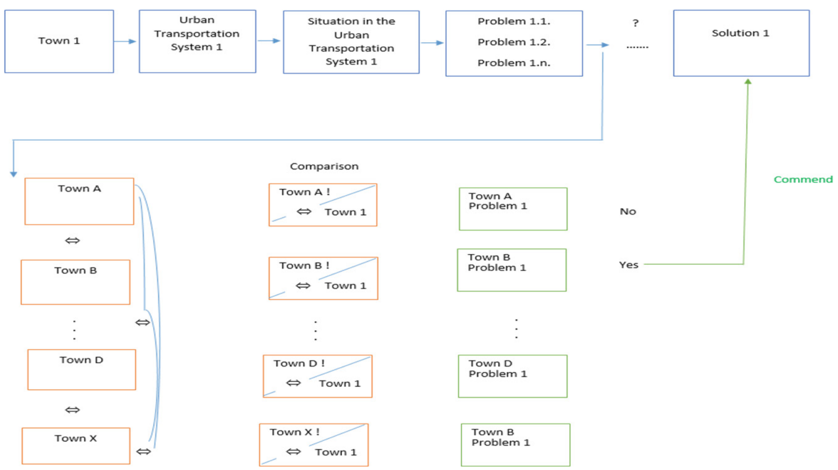

Fig. 1. Algorithm of decision-making process’ method

At the first stage of the research, it is necessary to consider the system of the city, to which range it include according to the Working Paper written by Lewis Dijkstra and Hugo Poelman, European Commission Directorate-General for Regional and Urban Policy (DG REGIO) [6]:

1. Densely populated area: (alternative name: cities);
2. Intermediate density area (alternative name: towns and suburbs);
3. Thinly populated area (alternative name: rural area).

It is necessary for research, comparison and situations' analysis of in the same cities in order to implement the experience of managing EU urban transportation systems (Table 2).

It is also necessary to review the urban transport system and the situation in the urban transport system. Then it is important to identify the problems that need to be solved in the urban transport system under the study.

Example. We want to implement in Ostroh the Estonian small towns' experience of managing urban transportation systems (see example Table 2).

Table 2. Data for comparison towns according EU

Towns	Criteria				
	Inhabitants (10000-19999)	Density (at least 300 inhabitants)	A minimum population of 5000	Less than 50% of the population living in rural grid cells	Less than 50% living in a high-density cluster
Ostroh (Ukraine)	15700	1436	+	+	+
Estonian towns:					
Haapsalu	11270	1097	+	+	+
Rakvere	15413	1629,4	+	+	+
Viljandi	17525	1199			
Valga	15044	842,2			
Sillamäe	13964	1307,5	+	+	+
Kuressaare	13276	890	+	+	+
Keila	10 012	889,9	+	+	+
Maardu	16 466	733,2	+	+	+
Võru	12965	1 108,4	+	+	+
Jõhvi	10051	1 319	+	+	+

At the next stage, it is necessary to assess how the investigated town is similar with cities to be compared. For such purpose, we form claims' system based on the study of cities. Listed claims' system is to be filled to the table for both towns. Then we compare the approval of the first list with the second list on this basis. We form a table to assess the general statements of the city as a structure, transport structure and management system (Fig. 2).

Claims' Town 1	Claims' Town A
1. Small town. →	1. Small town.
2. Minimum population 5000.	2. Minimum population 5000.
3. //////////////////	3. //////////////////
4. //////////////////	4. //////////////////
....	....
n. Local council's manager do not calculate the number of passengers who go off in each bus station.	n. Local council's manager calculate the number of passengers who go off in each bus station.

Fig. 2. Determination Towns' Equivalent Claims.

Based on comparisons of the list of claims of town, we propose to form a table of similar statements in order to estimate the Lorentz coefficient. Statements that describe town 1 (Ostrog), equal statements between systems, and statements that describe town A (Valga) necessary to write in the table (see Table 3). I case, when all statements are written down, it is necessary to calculate the similarity coefficient. This will allow us to look at how much the systems (towns or cities) are similar, which of the statements should be applied in the decision-making process.

Table 3. Towns' statements.

Town 1 (Ostroh) describes the claims	Equivalent claims			Town A (Valga) describes the claims
	The wording from the first	The wording from the second	The wording for both	
1.	Small town	Small town	Small town	
2.	Minimum population 5000	Minimum population 5000	Population	
3.	Density at least 300 inhabitants	Density at least 300 inhabitants	Density	
4.	Less than 50% lives in high-density clusters	Less than 50% lives in high-density clusters	Concentration	
5.	Intermediate area (towns and suburbs)	Intermediate area (towns and suburbs)	Intermediate area (towns and suburbs)	
n. Local council's manager do not calculate the number of passengers who go off in each bus station				
n + 1				Local council's manager do not calculate the number of passengers who go off in each bus station

The next stage is the construction of towns' similitude matrix, which allow us to compare the similarities of all cities among themselves (Table 4). It should be noted that the coefficient should not be less than 0.5 and the closer the value of the coefficient to 1, the more similar systems (towns, cities).

Table 4. Towns' similitude matrix

	Ostroh	Haapsalu	Rakvere	Viljandi	Valga	Sillamäe	Kuressaare	Keila	Maardu	võru	Jõhvi
Ostroh	1										
Haapsalu	1	1									
Rakvere	1	1	1								
Viljandi	1	1	1	1							
Valga	1	1	1	1	1						
Sillamäe	1	1	1	1	1	1					
Kuressaare	1	1	1	1	1	1	1				
Keila	1	1	1	1	1	1	1	1			
Maardu	1	1	1	1	1	1	1	1	1		
Võru	1	1	1	1	1	1	1	1	1	1	
Jõhvi	1	1	1	1	1	1	1	1	1	1	1

According to the data of towns' similitude matrix, we can see that towns are similar (coefficient similarity equal 1) and we can analyze the urban transport system and the situation in the urban transport system in Estonian towns. If the coefficient similarity lays in the range of 0.5–0.8 during the comparison of systems (cities), it is necessary to make a comparison of the general part of the common parts (Fig. 3). This will allow us to show detailed information about the similarity of cities and common towns' features.

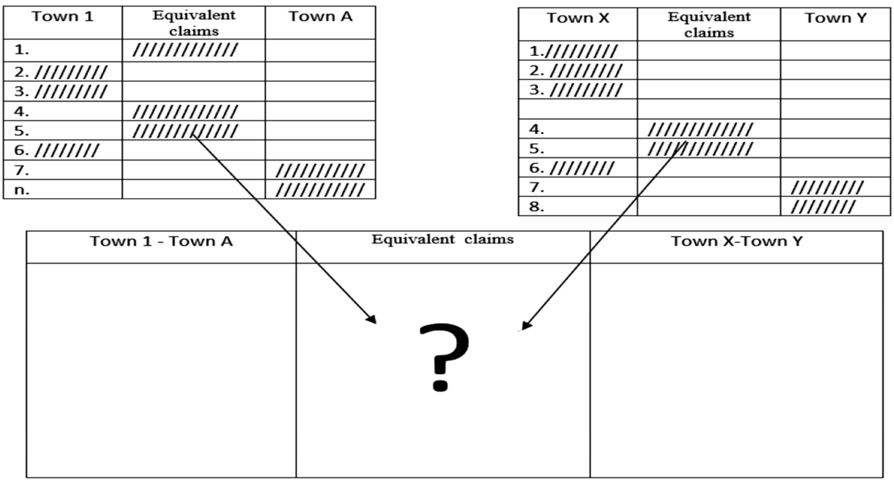

Fig. 3. General part of common parts

The next stage involves an end-to-end comparison of the statements of the city under investigation with another cities simultaneously. This will allow us to show the specific features of the systems under study (Fig. 4). It is also necessary to perform this, since during the comparison of the general part of the common parts of the cities, it may not be possible for the generality to be the same.

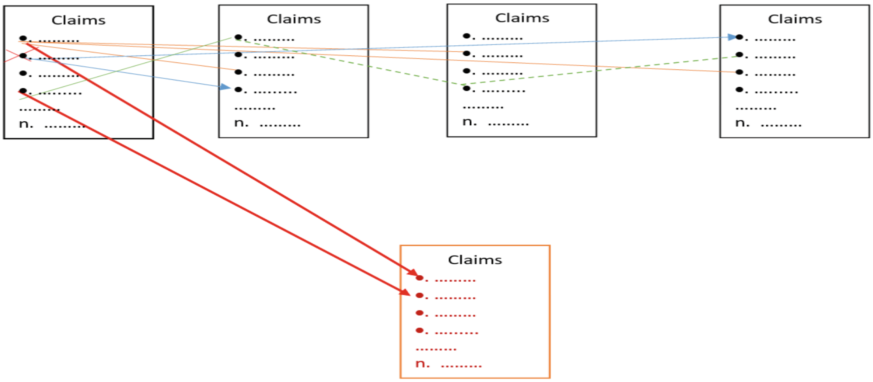

Fig. 4. An end-to-end comparison of the statements of all systems studied.

In order to form the final list of statements after such a comparison (Table 5), we will receive information about differences' features, which are in the urban transport systems and the management of the transport systems. Then it is possible to decide, which a positive experience in the management of public transport systems in Estonian small towns can be implemented for the town (Ostrog, Ukraine).

According to the results of the end-to-end comparison of the statements of all systems studied, we can see 19 similar features. The differences' features between Ostroh and 10 Estonian towns include impossibility to control the buses' time schedule, inconvenient public transportation, deficiency the buses' time schedule at the bus stations. Private carriers dictate their own terms and conditions for the provision of transport services, which leads to non-observance of the schedules on certain routes, deficiency an e-ticket, residents are forced to turn to the services of owners of private vehicles (taxi).

At the next stage of our investigation, we recommend to involve an end-to-end comparison of the statements between Estonian small towns. It is help to find out the similarities and differences' features in order to receive claims for decision-making process for Ostroh local council.

Table 5. Equivalent claims 11 towns (Ostroh-10 Estonian towns)

	Equivalent claims
1.	Small town
2.	Population
3.	Density
4.	Cells
5.	Concentration
6.	Intermediate area
7.	Classification
S.	Town-forming enterprises
9.	Diesel engine
10.	Thinking about ecological transportation
11.	Transport's line
12.	Frequency per month
13.	Peat hours
14.	Network of routs
15.	Transports' lines are important
16.	Permit for transportation after competition
17.	Frequency per route
18.	Calculation emission level is absent
19.	Control

3 Conclusion

Relying on similarity is one of the common ways of shaping decisions. Especially when there are just a few cases to deal with, but decisions have to be made. Whether it is good or bad, is not to be decided by us. However, it is possible to examine what its nature is and how it is implemented. Including if we want to implement or avoid implementing of the public transport arrangement in one small town that has proved useful in another city. So, in this work, we have somewhat explained some of the manifestations of similarity, including the descriptive similarity and the way it is numerally assessed. It was shown the application of the descriptive similarity as basic of the method possibility study and implement the experience of managing urban public transportation. We will present the results of the urban transportation systems similarity and statements to be applied in the decision-making process for Ostroh local council in the next paper.

References

1. Cohn, P.M.: Universal Algebra. Springer, Dordrecht (1981). https://doi.org/10.1007/978-94-009-8399-1
2. Jaccard, P.: Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. Bull. Société Vaud. Sci. Nat. **37**, 241–272 (1901)
3. Jaccard, P.: Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bull. Société Vaud. Sci. Nat. **37**, 547–579 (1901)
4. Lorents, P., Matsak, E.: Applying time-dependent algebraic systems for describing situations. In: 2011 IEEE Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA 2011), Miami Beach, FL, pp. 25–31. IEEE Press (2011). (IEEE Catalog Number: CFP11COH-CDR)
5. Lorents, P., Matsak, E., Kuuseok, A., Harik, D.: Assessing the similarity of situations and developments by using metrics. In: Czarnowski, I., Howlett, R.J., Jain, L.C. (eds.) IDT 2017, Part I. SIST, vol. 72, pp. 184–196. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-59421-7_17
6. Dijkstra, L., Poelman, H.: Regional working paper: a harmonized definition of cities and rural areas: the new degree of urbanization, p. 28 (2014)
7. Maltsev, A.I.: Algebraic systems. Science, Moscow (1970)
8. Szűts, I., András, B.: Decision support systems in management. In: 5th International Conference on Management, Enterprise and Benchmarking, 1–2 June 2007