KAUNAS UNIVERSITY OF TECHNOLOGY

MARIUS IVAŠKEVIČIUS

INFLUENCE OF URBAN SHAPE (AS MEMORY) ON SOCIAL CAPITAL

Doctoral dissertation
Humanities, History and Theory of Arts (H 003)

Kaunas, 2021

This doctoral dissertation was prepared at Kaunas University of Technology, Faculty of Civil Engineering and Architecture, Civil Engineering and Architecture Competences Centre, during the period of 2016–2021.

Scientific Supervisor
Prof. Dr. Kęstutis ZALECKIS (Kaunas University of Technology, Humanities, History and Theory of Arts, H 003).

Edited by Brigita Brasienė (Publishing House "Technologija")

Dissertation Defence Board of History and Theory of Arts Science Field:
Prof. Dr. Jūratė KAMIČAITYTĖ (Kaunas University of Technology, Humanities, History and Theory of Arts, H 003) – **chairperson**;
Assoc. Prof. Dr. Indrė GRAŽULEVIČIŪTĖ-VILENIŠKĖ (Kaunas University of Technology, Humanities, History and Theory of Arts, H 003);
Assoc. Prof. Dr. Tomas GRUNSKIS (Vilnius Art Academy, Humanities, History and Theory of Arts, H 003);
Assoc. Prof. Dr. Igor NIKOLIC (Delft University of Technology, Netherlands, Technological Sciences, Informatic Engineering, T 007);
Chief Researcher Dr. Renee PUUSEPP (Estonian Academy of Arts, Humanities, History and Theory of Arts, H 003).

The official defence of the dissertation will be held at 1 p.m. on 17 December, 2021 at the public meeting of the Dissertation Defence Board of History and Theory of Arts Science Field in Rectorate Hall at Kaunas University of Technology.

Address: Donelaičio 73-402, Kaunas LT-44249, Lithuania.
Phone (+370) 37 30 00 42; fax (+370) 37 32 41 44; email doktorantura@ktu.lt

The doctoral dissertation was sent on 17 th of November, 2021.
The doctoral dissertation is available on the internet at http://ktu.edu and at the library of Kaunas University of Technology (Donelaičio Str. 20, Kaunas LT-44239, Lithuania).

KAUNO TECHNOLOGIJOS UNIVERSITETAS

MARIUS IVAŠKEVIČIUS

URBANISTINĖS FORMOS KAIP ATMINTIES TALPYKLOS ĮTAKA SOCIALIAM KAPITALUI

Daktaro disertacija
Humanitariniai mokslai, menotyra (H 003)

Kaunas, 2021

Disertacija rengta 2016–2021 metais Kauno technologijos universiteto Statybos ir architektūros fakultete, Statybos ir architektūros kompetencijų centre.

Mokslinis vadovas
prof. dr. Kęstutis ZALECKIS (Kauno technologijos universitetas, humanitariniai mokslai, menotyra, H 003).

Redagavo Violeta Meiliūnaitė (leidykla „Technologija“)

Menotyros mokslo krypties disertacijos gynimo taryba:
prof. dr. Jūratė KAMIČAITYTĖ (Kauno technologijos universitetas, humanitariniai mokslai, menotyra, H 003) – **pirmininkė**;
doc. dr Indrė GRAŽULEVIČIŪTĖ-VILENIŠKĖ (Kauno technologijos universitetas, humanitariniai mokslai, menotyra, H 003);
doc. dr. Tomas GRUNSKIS (Vilniaus dailės akademija, humanitariniai mokslai, menotyra, H 003);
doc. dr. Igor NIKOLIC (Delfto technologijos universitetas, Nyderlandai, technologijų mokslai, informatikos inžinerija, T 007);
vyr. m. darb., dr. Renee PUUSEPP (Estijos dailės akademija, humanitariniai mokslai, menotyra, H 003).

Disertacija bus ginama viešame menotyros mokslo krypties disertacijos gynimo tarybos posėdyje 2021 m. gruodžio 17 d. 13 val. Kauno technologijos universiteto Rektorato salėje.

Adresas: K. Donelaičio g. 73-402, Kaunas LT-44249, Lietuva.
Tel. (+370) 37 300 042; faks. (+370) 37 324 144; el. paštas doktorantura@ktu.lt

Disertacija išsiųsta 2021 m. lapkričio 17 d.
Su disertacija galima susipažinti interneto svetainėje http://ktu.edu ir Kauno technologijos universiteto bibliotekoje (K. Donelaičio g. 20, Kaunas LT-44239).

CONTENTS

TABLE OF FIGURES

TERMS AND ABBREVIATIONS

ANN – artificial neural network, a broad term encompassing any type of computing inspired by the neuroscience, including but not limited to the perceptron's deep ANNs and capsule networks.

Biofilia – hypothesis that humans have evolved in nature and feel better in the natural environment or artificial environment with properties that mimic natural environment (Wilson, 1984).

Big data – refers to the amount of data that is too big to store, load or process in all at once in one computer. State of art machine learning methods usually depend on such amounts of data and utilise batch processing and cloud computing techniques to deal with it.

Breath first search – a strategy of exploring a mathematical graph by first visiting the closest nodes. An opposite strategy is the depth first search where algorithm tries to reach the farthest node as soon as possible.

Back propagation – algorithm that allows training of multilayer neural networks by allowing to derive adjustments to the biases of the neurons from the error function or adjustments to the biases of the previous layer of the neural network (Rumelhart et al., 1986).

Cellular automaton – is a useful simulation tool based on the grid. This grid is usually 2-dimensional; however, there are implementations of multidimensional cellular automata. Moreover, the grid is usually square; nevertheless, there are implementations for the most of the known grid types (Bays, 2012).

Coefficient of determination – as well known as R-squared written as R2 or R^2 method to calculate the prediction errors, tends to be used quite often. Its popularity could be explained by the fact that it is a comparison of two models in itself. It works by comparing the errors of prediction to the average of the target values:

$$R^2 = 1 - \frac{\sum(y - y_p)^2}{\sum(y - y_a)^2};$$

where y is the target value, y_p is the predicted value of a model, y_a is the average of all target values.

By comparing the calculated R-squared values of the tested models, it is possible to compare the models between themselves and measure the accuracy of the data model.

Corpus – sometimes, the text corpus in linguistics is a large and structured set of texts, used to perform statistical analysis on the text data, in the field known as Natural Language Processing.

Deep Artificial Neural Network – more commonly is shortened to Deep Learning, is a machine learning model that is taking a lot of inspiration from and giving insight into the studies of biological neural networks (Dechter, 1986). It is based on the chosen function regression where layers mean the regression of previous function outputs. There are specialized layers and layer architectures for vision, language processing, recommender systems etc. invented.

DNA – Deoxyribonucleic acid, the molecule that holds information needed for the creation and maintenance of certain organisms, if the right conditions are provided.

Geopositioning – process of identification of geographic location, usually by means of assigning data to the geographical feature: point, polyline or polygon.

Graph – in the field of mathematics, it is called the graph theory; graphs are studied as models of relations. It is composed of nodes and links, which represent the connectivity of the nodes. Many problems can be adapted to the graph theory and analysed with its methods. It is applied in many fields, such as computer science, linguistics, physics and chemistry, social sciences, biology and urban planning.

Latent Dirichlet Allocation – as well known as LDA, is the dimensionality reduction method used, but not limited to the natural language processing, for example, to extract subjects of documents composed of Word Vectors (Pritchard et al., 2000).

LSTM – long short-term memory, a type of deep neural network with the ability to predict the states of its own layers. Usually, it is used with the time series data or text.

Principal Components Analysis – shortened as PCA, a statistical procedure of dimensional reduction of the data matrix, which attempts to "rotate" the most correlated data points to be aligned, and then those points can be represented by the artificially generated dimensionless variable.

Power law – interesting propriety of power law distributions. If the plot is made where random positioned sphere radiuses were distributed by the power law, it is impossible to tell how much it is zoomed. When observing such an image on different scales, the distribution of details is the same (Clauset et al., 2009).

Voronoy cells – are regions, which divide the space where one region represents the shortest path to the point of interest. Function that produces Voronoy cells converts the points of interest to the borders of shortest path regions, stating that there is no preference to the points. It can be implemented on the cellular automata, but pure mathematical implementations exist as well (Voronoi, 1908).

Vectorisation – as well known as "image tracing", is the conversion of raster data to vector data, usually tracing boundaries of the same or similar value pixels (Itoh, Ohno, 1993).

Word Vectors – linear model of certain language, where words are represented by multidimensional vectors. There can be and usually are several models of a certain language with different number of vectors; the most common numbers are 50 and 100. When performing the principal component analysis on it, it is possible to get a map of words where close meanings cluster together. Probably, the most intriguing feature is that the word vectors enable arithmetic with words. It can provide a solution to following equation "king" - "man" + "woman" = "queen". Word vectors are created with unsupervised learning by taking huge corpus of language then trying to predict randomly removed words from the sentence. Interestingly, this is enough to capture the meanings of words (Grave et al., 2018).

Raster Data – is data presented on 2 dimensional matrix or lattice grid. It is compatible with cellular automata; therefore, certain types of cellular automata can be launched on the raster data directly. When used in Geographical Information Systems, raster grid coordinates represent the real world coordinates; in consequence, the grid size can be scaled easy to the interpret values (Bach et al., 1997).

RNA – Ribonucleic acid, similar molecule to DNR, usually used for temporary storage of genetic information.

Statistics – can mean numbers representing the results of statistical methods or a science that develops these methods. The purpose of statistics as science is to develop methodologies that allow testing hypothesis about the recorded facts (Romeijn, 2017; McIntosh, 2013).

Social Capital – there are many definitions of Social Capital; nevertheless, in this desertion, the definition by R. Putnam has been chosen: "physical capital refers to the physical objects and human capital refers to the properties of individuals, social capital refers to connections among individuals social networks and the norms of reciprocity and trustworthiness that arise from them. In that sense social capital is closely related to what some have called civic virtue" (Putnam, 2001a).

Shape Skeleton – representation of polygonal shape with tree like graph structure that starts from the corners of the original shape, continuing inwards as median axis. Such representation is suggested as a model of human vision and has statistical experimental evidence that supports it. Moreover, it is useful because it allows utilizing the graph analysis methods for polygonal shapes.

Shape as Memory – architectural theory where shape is described by using mathematical set theory, consisting of a set of transformations that are needed to execute in order to transform the basic geometrical shape (for example, circle) to its final complex shape. The author of the theory Michael Leyton claims that the maximisation of the number of transformations will lead to a new architectural style (Leyton, 2006).

Urbanism – unites multiple disciplines related with cities, either to plan, design or study them in various aspects, such as historical, social or economic. Such loose definition encompasses an abundance of Urbanism types, which are paradigms that are concentrated in specific aspects. The example types include: Networked Urbanism, Parametric Urbanism, Landscape Urbanism and many more in the collection of 60 Urbanism definitions (Barnett, 2011).

Urban Form – physical or planed state of the city, including all assets and connections. Land uses and densities as well as population density and diversity are parts of urban form (Anderson et al., 1996).

Urban Shape as Memory – is the adaptation of Michael Leyton's theory "shape as memory" to urban shape by substituting mathematical set theory with mathematical graph theory. This modification is based on the theoretical insights in the fields of psychology and neuroscience as well as experimental evidence. It is discussed in details in subchapter 1.3. The essence of this theory is that the shapes can be analysed by reconstructing the sequence of transformations that were used to create it; therefore, it can be adapted to the urban form as well.

XGBoost – machine learning algorithm that combines random trees and boosting (creating models from the errors of previous models) to make predictions. It is popular in the data analysis competitions.

INTRODUCTION

Background and motivation

Sustainable development consists of main three domains: environmental, economic, cultural and social. However, the social domain is the most over looked out of the four. The social phenomena, by contrast, are measured only by very negative indicators: crime, the need for social housing, the amount of benefits paid etc. One might think that the social system is approached only as a system, always in an extreme, crisis regime. However, this system is generally not in the extreme regime, there are just not enough sensitive indicators, and the tools to measure and influence them are simply not used. Therefore, the aim of this study is to pay attention to the social environment of cities, apply world-known indicators to Lithuania and look for measures on how these social indicators can be influenced in the urban planning. To this day, the social indicators are discussed with an emphasis on, perhaps, one of the most important social capital. There are two types of social capital, i.e., personal and local. This research is concentrated on the local social capital, which describes the shared values, social norms and trust of the location, which leads to the reciprocity and cooperation. Personal social capital is the amount and strength of personal connections or the ability to benefit through them. While personal capital can often be described as a negative thing, the social capital of the area is more positive. The very first use of the term social capital is to describe a place in an event of certain social initiatives that are becoming possible due to the increase of personal interconnections that are forming a society. As part of the theories of environmental determinism, the social capital could be partly influenced by the physical proprieties of the urban space. The most promising theory is to treat city as an extension of social consciousness or memory. It has theoretical foundations in behaviour psychology and neuroscience and is enabled by the data science and the emerging availability of computational power. It has derived from the theory shape as memory, which tries to model human cognition of shapes. Although there are many theories, which allow to model phenomenon concurrent in the cities: for example: space syntax, isovist, cellular automata, GIS and more; whereas in the process of city planning, it is inertly only static, and there are used only a few tools that allow predicting the social capital index and none that allow to predict the social index from the physical environment. Although the advances in data science and remote sensing resulted in “smart city” phenomenon, which slowly moves urban planning away from the stagnated spatiality, still, there are only a handful of examples of partial implementation of this philosophy, and there are no examples of a complete smart city with fully implemented philosophy. Further on, there are numerous methods to measure various parameters of urban shape as memory, which concentrate on the transportation or the type of usage, contrary to the methods of measuring architecture, where concentration is in harmony with visual information; therefore, the adaptation of methods that are meant for architecture to measure the urban shape as memory could lead to better models that would allow predicting social capital index.

Aim of the dissertation

The aim of this dissertation is to indicate the connection and regularities that were observed in the form as memory influencing social capital and propose a methodology using the Skeletonization of shapes and mathematical graph analysis as well as machine learning for the evaluation of the impact of urban shape on the social capital.

Objectives of the dissertation

In order to achieve the aim of the dissertation, several smaller objectives must be achieved:

- To analyse the existing methods and theories that help to support the concept of “urban shape as memory”;
- To analyse Leyton’s methodology "form as memory" and adapt it to the urban form by using straight skeletons;
- To investigate how big data can be used in the proposed straight skeleton analysis methodology;
- To adapt the methodology that it would be compatible with the state of art Artificial Intelligence, Machine learning models;
- To evaluate several machine learning models with the coefficient of determination in order to select the best suited for the prediction of social capital index; suitable social capital index data is available in the USA;
- To produce the measurement of urban form as memory and test if modelling is reproducible in the second selected location (Lithuania);
- To explain the connection regularities of urban space as memory to the social capital.

Statements for the defense

- Urban form as memory is statistically significantly related to the social capital.
- The theoretical concept of “urban shape as memory” can be practically measured by using the developed methodology of graph analysis of the skeleton shape of buildings and the space between the buildings.
- Shape Skeleton model is a good approximator for the comprehension of the surrounding environment in humans.
- The methodology to measure “social capital” in administrative units is adaptable locally and produces data that by itself is beneficial for understanding the tendencies of urbanism.

Scientific novelty

Although there are many theoretical implications that social factors are influenced by the physical environment, only a few small-scale, narrow field attempts were made to measure it. No studies were found that reported the modelling urban form influence on the social capital. Furthermore, no studies were found that measured the urban form as memory. This research offers an original methodology using the Skeletonization of shapes mathematical graph analysis and machine learning. Skeletonization is a process of finding the median axis that represents simplified topological structure of the shape. This methodology has not been used before to analyse the urban form and predict social statistics such as social capital.

Practical implications

Urban form as memory can offer insights into understanding urbanism and urban form, offering a new perspective. There are several methodologies designated for the urban form evaluation, but this one is exclusive, because it originated from the ideas that relate to the influence of comprehension of physical environment on the individual behaviour. Urban form as memory could be used to model social capital as well as other important, but not necessarily social indicators that are related to the urbanism, such as, but not limited to pedestrian traffic, crime prediction, perceived passage of time, location allocation and many others.

Modelling social capital from the result of urban form analysis could change urban planning in unprecedented way, because without it, it was impossible to predict social dimensions of the city; therefore, the urban planning for social capital optimisation could not exist. Such model could offer predictions of social capital index in the stage of city planning; therefore, the city plans could be made for the maximisation of social capital.

A list of the most influential urban forms as memory to the social capital properties could be made. It could offer insights and strategies for urban planners that are striving to maximise the social capital index in their city plans.

Although many models that offer various predictions influenced by urban form are developed, none is used in the city planning; therefore, this methodology could be added to the critical mass of urban methodology. After this critical mass is achieved, it is impossible to ignore, and the urban planning would rise to a more intelligent level.

Methodology

The dissertation is based on multiple methods. Literature analysis was used for inductive and deductive reasoning, which lead to the selection of practical methods of gathering and analysing data. Urban form as memory is a new multidisciplinary model of urbanism that is consisting of methodologies from the fields of psychology, mathematical graph analysis, neural science, simulation and architectural theory. The social capital index is measured with the methodology of the field of sociology, in which, the core is the statistical analysis methods, including the principal component analysis. Social capital and urban form as memory data set is spatial; therefore, it has to be unified by using Geographical Informational Systems. Finally, the modelling accuracy is measured while testing the machine learning models, such as linear regression, lasso, artificial neural nets, deep learning, LTSM and XGBoost. The data analysis models have to be validated with the validation data set, measuring the coefficient of determination. All methodology is tied closely together; separate methods that are used in the research could be classified as multidisciplinary; therefore, when used together, they form a transdisciplinary approach.

Structure of the dissertation

The dissertation consists of the following sections: introduction, literature review and analysis, methodology and results, discussion and conclusions, references and a list of publications.

The volume of the dissertation is 96 pages. The dissertation contains 24 figures, 2 tables and provides 285 references.

1. CITY IN A NUTSHELL: CONCEPTUAL APPROACH TO THE LITERATURE REVIEW AND ANALYSIS

In order to gain more fundamental knowledge, the fields surrounding urbanism will be explored as well as the studies that refer to the comprehension itself and the feelings it may induce. The feelings are motifs for action; thus certain behaviours, when considering all possible behaviours, some of which are of most interest for the purpose of this dissertation, are social. Different theories exist in the literature regarding how they analyse their subjects; theories “in a nutshell” resort to classification; however, the classification is used when measurements are impossible, but more refined theories offer methods of measurement, which allow more precise comparison of their subjects. The world is continuous, where most things can be measured, because by agreement, they have assigned measurement units. There are countless of measurement methods that were invented, giving answers about the properties of the world in their corresponding units. The weight in grams, distance in meters, time in hours, to name a few; however, there are measurements that are behind difficult ideas, for example: jerk is the speed of change of acceleration, or fractal dimension is the propriety of roughness of shape. Nevertheless, not all questions can be effectively answered by continuous measurements: for example, the space is free or occupied, the object exhibits certain proportion, or it does not. Therefore, the method “in a nutshell” of this dissertation will be used to classify the theories into two parts, where the term “canonical systems” will be used to refer to the theories and methods that use rules for the classification or quantisation into discrete parts, and “continuous systems” for theories that offer methodologies, which enable measurements of proprieties of subjects under the analysis.

The city can be analysed as complex systems, like ecology, economics, sociology and culture. Such systems are sustainable, in other words, they are resistant to change. Unfortunately, such sustainability is a two-edged sword, when resistance is overcome, the system transitions to an extreme less efficient mode, in which it as well resists to the return to the effective mode. Ecological catastrophes, economic crises are examples of the extreme regimes in complex systems. Complex systems that are stuck in such extreme modes do not return to normal mode for a long time and without effort, even after the conditions that caused mode return to their previous state; usually, it is not enough to switch the mode of a complex system back. Sometimes, it does not even depend on time, but it depends on the surrounding conditions, which need to improve significantly for the system to return to the efficient mode. Another propriety of a complex system is emergence, when several simple subsystems form another system, which is bigger and greater than the sum of its parts. The phenomenon of emergence has been observed in most fields that are mentioned in this paragraph, including the one, which is of the most interest for the research, i.e., urbanism. The basic idea of sustainable development, which originated in the economy and spread to the other areas, is to know the system well enough and use it responsibly that the stakeholders can keep it in a normal efficient mode without weighing it into an extreme. Ecology, economics and even culture have enough sensitive indicators to show the state of the system. These indicators can be used in the urban planning because the tools that can influence them are known and widely used.

One of most useful things that the science can achieve is to predict the future. In order to achieve this, one must record the conditions and outcomes of those conditions, either occurring naturally or as the setting and the result of an experiment. Only by looking at the recorded data historically, wise men were able to make useful insights and derive connections, which were written in mathematical language and became known as the laws of nature. Later on, this process of making insights was analysed and became the science of statistics. It allows to predict if there are relationships hidden in the data; moreover, it provides methodologies and algorithms that allow production of mathematical formulas. As most statistical calculations are not possible without a computer, the machine learning algorithms, which are based on statistics, are only possible and are enabled by computers. Although machine learning models are often too difficult for an average human to understand, and experts have to use special methodologies to interpret and explain their inner working, they produce predictions that are impossible with other methods. Some of the machine learning models, notably artificial neural nets, are inspired by neuroscience; moreover, besides the predictions of data, they offer insights about the working of biological brains. One could state that in the age of information and machine learning, everything can be modelled and predicted; however, it is not true, at least for now. In order to use the statistical analysis and machine learning methodologies of model construction, the data matrix should be produced where rows represent the observations and column initial conditions and result. There are a few exceptions in the image analysis, natural language processing and other fields, but in a nutshell, they have the process of producing data matrix automated. The essence of the data matrix production is the measurement or evaluations of other data types to provide meaningful numbers, representing the phenomena that someone wants to analyse. In order to be able to produce models that can predict social capital by analysing the urban shape as memory, they should be measured or described with numbers or parameters. The history of human effort should be reviewed in order to measure or model environment or its features related to the cognition of environment, which is essential to the urban shape as memory in various ways. Although aesthetics and navigation from first glance could be separate subjects; in the scope of urban shape as memory, they are all related to the cognition of environment. Therefore, by studying various fields related to the cognition environment, it would be possible to gain insight about how to use it in research, despite the fact that separate studies focus on the aesthetics or navigations. The term "urban shape as memory" originates from Leyton's term "shape as memory", which states that every shape has history of its creation stored in it, and theoretically, this history could be extracted, and human brain does it; therefore, it uses this process as a comprehension of space, but no one was able to produce a working model of this process yet. There are three ways to think about the cognition of space: just philosophically without trying to find any applications, trying to guess the rules, which were used by nature or human to create the environment, and trying to find the methods of measurement that have relations to the cognition of environment. As cognition can be viewed as an emergent result of community of neurons in a biological brain, city can be viewed as an organism with a community of people that are forming emergent result of social capital. It is possible to try and find evidence in neurobiology or by experimenting with animals. Furthermore, it is possible to study artificial environment creation cultures and how certain direction was taken in the creation of the environment that inspires social interaction. They will be discussed in such order in the following chapters.

1.1. The defence of memory palace: efforts to rationalise human cognition of space

The existence of "the method of loci" or more popularly known as memory palace, the imagination technique that uses the imagined environment to enhance ability to memorize anything is the most pure proof that the cognition of environment is wired to the core of human memory (Yates, 2014). Urban space as memory refers the special state of urban space that only exists as the recording in memory, after it was processed by the cognitive procedure to the form, which is stored in memory in the human brain. There are studies which analyse the urban space from this perspective, most notable *The Image of the City* (Lynch, 1960). In essence, this theory states that an individual's memory of the city differs from reality, but some repetitive elements persist between the individuals. These include paths, edges, districts, nodes and landmarks. Paths are identifiable connections (streets, walkways, canals) through which individual could move. Edges are identifiable borders dividing several zones. Districts are those zones. Nodes are central places, usually preferable meeting or gathering locations; they usually, but not necessarily, have landmarks, which become a part of their identity. Landmarks, the exceptional natural or human created objects (sculptures, monuments etc.) not necessarily have to be in the nodes. Some combinations of these elements can have a synergetic effect; others can interfere with each other. The lack of elements creates a gap of identity, which distorts the mental image of the city. Although this theory has its own methodology, consisting of qualitative social surveys, it was modernised and adapted to the digital age with methods like isovist (Benedikt, 2016). There are many more; however, their approach to the mechanism of surrounding space comprehension could be analysed more thoroughly; for this reason, a deeper dive into the history of such attempts is necessary to analyse specialised and universal methods.

The feeling of beauty and harmony is probably inborn; evolutionary psychology states that it has evolved because it was beneficial to the survival of prehistoric ancestors (Dutton, 2009). However, it is very hard to explain, measure or quantify. From ancient times, the thinkers have tried to make sense of this feeling and explain the conditions when it is triggered. Ideally, it would be an analytical method, resulting in the continuous value, representing an accurate amount of potential feeling that was initiated. Nevertheless, the simplest approach is to guess or invent a rule, which could explain the process; then, while using it in the creation process, it would assist in achieving the aesthetical value.

1.1.1. To leave or not to leave: the historical models of comprehension of physical surroundings

Humans, as a part of the animal kingdom, evolved in vision in such way that it is most useful to recognise simple things: food, safe and dangerous situations as well as mating partners. However, pure geometrical forms are relatively recent, e.g., Egypt pyramids and later creations. This is compatible with "triune brain" theory, which divides the human brain into three parts: reptilian complex, limbic system and neocortex (MacLean, 1990). Reptilian complex is very fast, but can make only very simple functions like aggression, dominance, territoriality. The limbic system is responsible for the motivation and emotion involved in feeding, reproductive behaviour and parental behaviour. Neocortex gives mammals the ability of language, abstraction, planning and perception. The human visual nerve is connected to the Lateral geniculate nucleus, which is a part of Talamus, and in the Triune brain theory, it is in the reptilian complex (Cudeiro, Sillito, 2006). Therefore, it is only logical that the first who is going to react is reptilian complex, later the limbic system can react to override the reaction of the reptilian complex, and further on, the same repeats with the neocortex.

The ability to recognise safe and dangerous environment must have been essential for the survival of prehistoric people; therefore, the ability to classify them and output a corresponding feeling emerged in their brains in the process of evolution (Dutton, 2005). In the course of the history, this ability to classify the environments allowed to develop skills for the modification of environment with the purpose of influencing this classification and inducing a corresponding feeling. The functional classification of architecture is compatible with this approach: temples invite to relax, meditate and listen to the words of wisdom, inducing feelings of safety and harmony, grotesque proportions that allow a good defence of the castles as well induce the feelings of unapproachability and fear in enemies, encouraging them to leave. Having inborn feeling emerging from the subconscious environmental classification allowed to notice the properties of the object that induces this feeling; moreover, the collection of such properties became a collection of rules or canons used to modify the surroundings with the purpose of inducing such feeling. The modification of surroundings in purpose of achieving aesthetical feeling started in prehistoric times as canonical art; the most famous for its cannons is ancient Egyptian art (Davis, 1992). Canonical systems were used for city planning as well; this approach is thoroughly analysed in Lynch's theory of "good city form" (Lynch, 1984). Starting with the already mentioned Ancient Egypt with the example city Kahun 3000 AD, where the separation of citizen classes is clearly expressed, it continued to the Roman empire with the example city Timgrad 110 AD, which was built following the rules of building military camp, which are described in the *De Architectura* that is regarded as the first book on the architectural theory (Vitruvius, 15 ADa). Alternatively, the urban canons in China regulated the dislocation of certain functions and their aesthetics as well as the materials and colours, explained by the magical beliefs. Further on, in India, the urban city model was based on the belief that city has a spirit and was planned using a picture of "pinned down demon", where certain parts of its body resulted in the type of use of location and requirements. All the mentioned ancient canonical systems are surmised in the city as organism with different "parts working together", comparing city centre to the heart that pumps traffic and businessperson, intellectuals to the brain of the city, neighbourhoods to the city cells, urban sprawl to the infectious spread (Lynch, 1984). Synonymous to organism could be an organic city or organic growth of the city; for some researchers, it only means that it was unregulated or poorly regulated, without planning (Kostof, 1993). Others argue that even planed cities sometimes exhibit patterns of organic growth (Michael et al., 1994a).

Artificially amplified feelings demanded explanations; therefore, the canonical analysis was invented. Firstly, the golden ratio is perhaps the most famous and oldest. It was described by the ancient Greeks for the first time (Euclid, 300 AD). Another one, which produces a sequence of numbers, which have golden proportions to the closest members of increasing accuracy, Fibonacci sequence was first described around the same time in India (Pingala, 200 AD), but in Europe, it was first described in 1202, in the same revolutionary book that initiates the change of numerical system from Roman to Arabic (Euclid, 300 AD). Fibonacci was describing this sequence as a means to predict growing rabbit under the supposed ideal conditions. However, the connection between two was noticed much later in Germany (Jacob, 1594). In the Roman empire, the golden ratio was used as a recommendation to be used for proportions in architecture, and the human body was the main example that possesses this proportion in several cases (Vitruvius, 15ADb). The famous illustration by Leonardo da Vinci *Vitruvian Man* is inspired by the proportional insights. The same principles were continually used as architectural tradition and popularised by many architects, including Corbusier (Corbusier, 1923), and in modern times, Neufert (Neufert, Neufert, 1970) and Ching (Ching, 2007). Golden ratio and Fibonacci sequence are measured or derived values of imaginative ideal philosophical concepts; therefore, it can be easily applied to the canonical system as in the *De Architectura* by Vitruvius, but it is more difficult to use as a measurement for aesthetics. Nevertheless, Golden ratio was used for the analysis; however, the process is not straightforward and requires expertise; therefore, it is a subject of inaccuracy based on the personal preference (Markowsky, 1992). The golden section rule is the most known design rule, but the method of application is not very clear and requires human skills to be applicable. However, the usage is controversial, it could be read in the centre, but divide the piece in golden section proportions, and there are works that use it in this way, and achieve the aesthetic result. Nevertheless, there are plenty of art examples that place symmetry axis in the centre (therefore, defy the rule) and are still considered masterpieces or even benchmark of golden section usage. For example, Salvador Dali's *The Sacrament of the Last Supper*. Some human and nature creations exhibit a noticeable number of rules that are being used to create them in the patterns containing self-similarities. By analysing the patterns in several cases, it was possible to reconstruct the rules of the creation process. The rule reconstruction examples mostly follow the same process: there is the starting state, then the rules are applied, and the final state is reached, which is the resulting object. Such approach can be applied to architecture, urban design and other fields as well. Moreover, the golden section rule does not allow to measure how aesthetic the object of the analysis is, because it is not known how many times one must use the rule to make the object aesthetic. The question is what happens if the rule is used only on the parts of the object. Facing uncertainty in the unanswerable questions, the countering methods were developed that do not require human expertise or interpretation.

Such systems that require human expertise, despite their shortcomings, are developed until now. For example, in the theory under development "Periodic Table of Architecture", sixteen adjacency philosophical perspectives are described. By reviewing them for a certain subject, the act of building insertion into the existing environment can be evaluated in a very sensitive way. This methodology adapts to different urban environments and can be readily used for evaluating free standing buildings in nature. However, it requires very high degree of expertise and deep understanding of those sixteen concepts, but the masses will be pointed out later, struggling to differ between simple concepts as beauty and professionalism (McKay, 2019).

Another modern example could be the one related to biophilia, a concept of healthy physical proportions that will be discussed widely in the following chapters. There are ten evaluations of architecture concepts that are presented, while some of them are simple, e.g., is there water (ponds, fountains), used in architectural subject, others are difficult to explain and evaluate, e.g., "Organized-complexity: intricate yet coherent designs — and extends to symmetries of abstract face-like structures". Using the methodology, the evaluator should choose a mark for each concept: 0 if not present, 1 for some, 2 for abundant. The sum of all marks is biophilic index that is showing how healthy is the influence of a building on its users (Salingaros, 2020). Although the methodology is quite clear and most concepts of the evaluation are backed up by the experimental research, it is lacking examples of execution.

The canonical analysis systems show an attempt to understand how physical environments induce feelings in people and consequently their behaviour. They have simple principle: “if this, then that”, which is easy to understand in many cases and is the essence of classification, dividing the world into granular fragments. It represents the multimodal operation of the human mind as a complex system, the natural tendency of inventing classes or more popularly known as stereotypes. Although canonical systems do not offer methods that could be used in this research, because they do not offer methods of measurement or evaluation; nevertheless, they lay down the theoretical foundation, on which other more useful methods were developed, which will be analysed in the following chapters.

1.1.2. More canons are needed: invention and uses of the shape grammar

Canonical systems are fundamental in understanding the process of creation; therefore, they have to be reviewed thoughtfully. Words like elegance and beauty have been used to describe equations in mathematics (Hardy, 1921). There is a special feeling that is called "aha moment" when you start to understand, or to fit the mathematical idea into one thought. There could be a shorten explanation of the same idea to one paragraph or one long sentence, but then it only represents this idea but not feeling. Therefore, the mathematical language compresses every variable to one letter and every function to one symbol, if possible; thus, in order to read it, one must know the hidden meaning of every letter and symbol. However, when one does by looking at the equation, the idea is seen as well as the feeling. Therefore, the capturing of the feeling of beauty or elegance is achieved by the maximisation of information compression.

Although the field of mathematics is very comprehensive and has many interesting aspects, for the sake of this argument, the discussion will be focused only on the compression of information. Describing the design rules and application in plain text can expand to several pages and could be difficult to comprehend. It has already been discussed how mathematical language can describe the laws of nature to the equations that are containing just several symbols. The same approach can be adapted to the canonical designs, and this is known as a branch of applied mathematics, named formal grammar, first described by Chomsky (Chomsky, 1956). In the formal grammar, the artificial, specialised language is created where strings of characters are composed of a specialised alphabet and according to the languages syntax (it should not be confused with the space syntax, which sounds similar, but is from a completely different field; it is known that the space syntax is created as an argument to shape the grammar inflexibility; thus, some terminology seems to be borrowed). The alphabet determines a list of legal characters in certain application, and syntax provides methods how the string of the characters can be transformed. It is counter intuitive how syntax does not offer an explanation of the characters, because it is not a part of the formal grammar. Therefore, formal grammar can be used to generate strings of characters, which later can be interpreted by previously known rules; the results can be found in the spoken language as well. Hence, the formal grammar is sometimes referred as the language generator. This definition is quite broad, but it can be concentrated to make specialised applications. Formal grammar theory is quite universal, it is applied in computer science, linguistics, biology; however, the connections to visual environment are mostly related to this research. One relevant usage of this system is to assign the expressed genes to the DNA sequences (Datta, Mukhopadhyay, 2013). The decoded DNA molecule is expressed in a sequence of characters, representing modular parts of a molecule, called nucleotides. The processes of how enzymes are presented in the cell's nucleus to use it to assemble proteins can be analysed in term of the formal grammar.

Interestingly, this is not the first usage of the formal grammar; apparently, Panini used it to describe Sanskrit grammar in the 4th century BCE (Pullum, Kornai, 2003). Many interesting formal grammars are invented; surprisingly, there is even one dedicated to describing knots (Knight, Stiny, 2015). Formal grammar has a branch, which is related to this research. It is dedicated to describing shapes and is named shape grammar (Stiny, Gips, 1971). The shapes that are described in this paper have self-similar shapes that could be described as fractal, although the authors never use this term, despite the fact that the term "fractal" was already introduced (Mandelbrot, 1967). Stiny has made plenty contributions to the field of shape grammar, showing that it is universal enough and timeless to be applied to a wide range of cases. In China, there is a craftsman tradition to make various window lattice patterns (Dye, 1974); one of them, in which Stiny was interested in, is named "ice-ray". It imitates the cracks that are formed on ice in nature in a stylised way. Although the actual patterns have variations, they seem to follow the same shape grammar, which was reverse engineered and described (Stiny, 1977). His effort was continued by implementing "ice-ray" grammar to MATLAB plug-in for turtle graphics (Lee, Tiong, 2013). Moreover, an interesting research has been made on the evaluation of Palladian plans. In case of the research where the subject is extremely regular, it is possible to recreate it with the shape grammar. In the evaluation of Palladian plans, Stiny recreated the plans of historical Italian villas by using shape grammar and used the analysis of grammar to evaluate the aesthetics, which has derived from the unity and variety (Stiny, Mitchell, 1978). Specialised shape grammars were invented for the design of Mughul gardens (Stiny, Mitchell, 1980), Japanese tearooms (Knight, 1981), Hepplewhite-style chaif-backs (Knight, 1980). Later on, Stiny described the method that can be used to construct own shape grammars (Stiny, 1980). Shape grammars remain in use even in contemporary times to analyse the workings behind the ideas of modern (Flemming, 2016) or contemporary (Eilouti, 2019) building facades. Shape grammars were used to procedurally generate full 3D models of urban sprawl (Halatsch et al., 2008).

In the examples analysed in this chapter, the shape grammars are specialised to a very narrow and particular purpose, although they are interesting and valuable, some of them were even used for measurement. However, it would be impossible to use them to measure the urban form, because it has too many variations, which could not be covered by a specialised system. Nevertheless, the shape grammar analogy with DNR could be used in the extension of city as organisms: they share the propriety of being emergent phenomena, and emergence is a propriety of complex systems that are the core of this research. Therefore, the question is whether these methods could become the foundation of a universal shape grammar.

1.1.3. The gene of emergence: canonical systems which can create anything by using only simple rules

It has already been discussed how a city can be perceived as an organism; thus, the organisms generally exhibit self-similarity and roughness on numerous scales as well known as fractality. The cities were analysed as fractals, although this discussion will be provided in the following subchapters, in this subchapter, the universal formal grammar system, which possesses properties connected to the DNR and fractality, will be reviewed.

Fractality of nature, self- similar shapes that are efficient to produce by using emergent proprieties of a system that translates DNA into other molecules, could be mirrored back to the part of itself that has to be interpreted in the brain (Losa, 2014). Inaccuracies and overcompensations of the human vision system are known and abused by smart compression systems in computer graphics that deliberately loose information that is insignificant to the human vision (Sonka et al., 1993). Images that are produced by the compression algorithms can be evaluated by other algorithms winch model as the perceived error of human vision (Gore, Gupta, 2015). Some of the compression algorithms move even further and take advantage of self-similarities that are present in the pictures of nature to perform fractal image compression, which is based on the visual perception (Chen, 1995). Although fractal compression was invented earlier and showed better compression rates than the compressions that are commonly used today, its newer applications went mainstream because of higher computational costs (Barnsley, Sloan, 1990). However, the human brain, as a computer with fractal proprieties and critical operation modes that as well point to the fractality, could be more efficient in the fractal compression (Kitzbichler et al., 2009). It is supposed that at some stage of human vision, the brain performs fractal compression, and other systems are accustomed to the success of fractal compression. If such vision system is exposed to the image that induces ineffectiveness of fractal compression, such as plain walls, patterns without variation of scale, the rest of the system suffers from the information overload and results in the induced anxiety and all its consequences. Tailor named it "fractal fluency" in his research where he found that fractal images with mid-range fractal dimensionality reduce stress (Taylor, Spehar, 2016). This effect is measured when recipients observe Pollock paintings (Taylor et al., 2011). The measuring fractal dimension of genuine and fake Pollock paintings can be an accurate predictor of counterfeit works (Shamir, 2015). Although civilisation possesses a skill base, knowledge and natural background of beneficial fractal usage; sometimes, it is not enough to utilise it or transition it to the traditional behaviour, considering the cultural and political pressures. Still, there have been studies that continued to prove that the properties of artificially created environment, architecture and urban space are important to various aspects of human cognition, sense, health and sociability. The creations of nature possess fractality and scale distribution proprieties; therefore, it could be recreated by mimicking such process. Nature uses information coded in DNR molecule and emergent processes of inner cell structure to accomplish all varieties of life; consequently, the artificial system could use a string of symbols to code information and decoder module to interpret it for emergent result. Primarily, this system was created to simulate simple multicellular organisms; however, later, it was expanded to mimic the shapes of plans and fractals. Lindenmayer noticed the mathematical elegance in biological nature; therefore, he invented a mathematical method to describe the plants named L-System (Prusinkiewicz, Lindenmayer, 1990). This system is completely customisable at all levels. It is based on the idea of a computer language logo, created for teaching purposes, where simulated robot "turtle" is controlled with special commands (Abelson et al., 1974). It has the abilities to drive forward, backward, turn left and right at any angle and distance; it has a pen, which can be lifted up or down; therefore, when it is down and mowing, it draws lines. Some implementations show turtle with the ability to hide it; the problems, which students are trying to solve, lie in the images produced with the turtle's pen. L-System is easy to understand; some simple examples can be worked out in one's imagination or on paper; therefore, one example will be analysed to prove it.

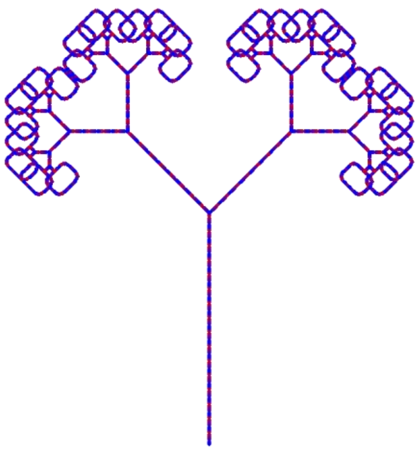


Figure 1. Pythagoras tree drawn with L-system (author's implementation)

The essence of L-system is a specific custom programming language, which is modified for every fractal. It consists of commands, represented by one character, which can be of two sorts of variables or constants, where variables are those that could be modified, and the constants remain unchanged. Then, there is an axiom, i.e., the primary state of the fractal, and rules that are applied for every iteration. L-System starts with an axiom and applies rules to it several times, specified by the user, to make longer and more complex; then, it executes it to produce the final shape. The properties ant process of construction of one fractal will be analysed: as just before this, there was a discussion about trees, it will be a binary tree. It is named binary, considering that the trunk and every branch have two branches. Lindenmayer noticed that the branches mostly start at a certain angle from the trunk, and later when the branch get older, they produce their own new branches, using the same angle, therefore, in this fractal standardised move, which can be persistent for certain shape simulation. It has two variables "1" and "0", meaning "draw line segment" and "draw line segment ending in leaf". Two constants "[" and "]", meaning remember or write to memory state, consisting of position and rotation as well rotate 45 degrees left, and recall from the memory or move to the recalled state as well rotate 45 degrees right. The memory refers to the stack type of memory, where it is impossible to choose an element from the memory; instead, the last written element is returned first and deleted upon return. The write and return are named "push" and "pop", accordingly. The axiom of this fractal is "0". It has two rules: for every iteration, every "1" is replaced with "11", and every "0" is replaced with "1[0]0". If simplified, the first rule makes segments without leaves twice longer, and the second rule replaces the segments with leaves into two branches with leaves directed 45 degrees away from the current axis. Therefore, it starts with "0", and after the first iteration, it becomes "1[0]0". It should be noticed that the first rule was not executed, because it is executed first, and at first, there were no "1" to be replaced. After another iteration, it becomes "11[1[0]0]1[0]0", and it starts to become difficult for a person to interpret, but computers do not have any problems even with 1000 times longer strings. As it is possible to see, it starts with trunk, which is two segments long, and then, it starts with the left branches and gradually passes to the right branches as it finishes. For every following iteration, straight branches consisting of "1" will become twice longer, and all branches with leafs will and produce twice as much branches of leaves. It has fractal dimension of 1.32, and the proportions of its parts follow the golden section rule. Visually, a binary tree is very similar to the straight skeletons, which will be introduced later. Formal grammar is an integral part of the L-system, the axiom is the primary string, and the rules are syntax in the terms of the formal grammar. However, L-system pushes the idea of using rules to transform the strings even further by allowing the rules to be used multiple times to achieve emergent results in a similar way the nature is using DNR sequences.

Many mathematical fractals were discovered before the L-system; however, they had their own unique and specialised algorithms that could generate them. L-system allows to replicate most of them and add natural fractals, plants, trees etc. With the concepts of city as organism and city as fractal, it is only natural that the next should be a city as the L-system. One could imagine a genotype of urban pattern coded in such a string, and when executed, it could produce an actual simulation of urban growth. The extension of L-systems, which allow global goals, and local constrains were developed to model a fabric of a city by separating the land with the transport system, later subdivided into land plots and even populating the buildings with height, resulting in a believable 3-dimensional model of a city (Parish, Müller, 2001). Another group of researchers used multiple algorithms for the generation of city structure, and an algorithm has derived from the L-systems to construct street networks within the blocks formed by roads. The algorithm had control parameters, which allowed it to be configured to generate different patterns of street networks, i.e., raster industrial and organic (Kelly, 2007).

It was mentioned earlier that the organism that used DNR to construct molecules can be described as formal grammar, but Lindermeyer does not mention inspiration or reference of his system to the DNR mechanics; therefore, it is safe to assume that this coincidence is accidental, not that it diminishes the discovery in some way. The process of coping one strand of DNA to RNA molecule is used to make any molecule that is used in the body by the process named translation, which takes place in ribosomes, where new molecules are made by matching the combinations of three nitrogenous bases, named triplet codons. This is analogous to the process in L-system where the shapes are translated from strings of characters and Leyton's process grammar where transformational matrices deform simple shapes. Interestingly, neither Lindenmayer nor Leyton discusses this parallel. In L-system, as in the real organisms, there are information sequences and rules. When the rules are followed to interpret the information sequence, the result of complex fractal shape is produced. L-System is an interesting model of life, because it can produce shapes that are recognisable as natural, and the emergent process of interpreting linear information produces multidimensional shapes and structures. Every living being is a result of the emergent process of interpretation of linear information that is coded in the molecular structure of DNA. Deoxyribonucleic acid (DNA) is a double stranded molecule, coiled like a spiral, named double helix. It is composed in a similar manner as a string of L-System that is made of characters; the parts of DNA are called nucleotides. Nucleotide of DNA has three components: five carbon sugars, a phosphate group, and a nitrogenous base. Sugar and phosphate are always the same, but there can be 4 different nitrogenous bases: cytosine, guanine , adenine and thymine, shortened as C,G,A,T. Double strands are needed to simplify the copying and error correction, which is made by attaching several proteins like helicase, primase and polymase; it is utilised in the cell division process called mitosis. This is done inside a cell nucleus naturally, but it can be reproduced in the device named thermo-cycler (Weier, Gray, 1988).

L-system universality comes from the coincidence that this model works in the same way that the actual biological systems work, i.e., they read a string of coded symbols and perform an action, according to the meaning of a symbol. As L-system can model many natural and artificial patterns, including patterns of urban form, it could be considered universal. L-system is important because it is canonical and crosses boundary, which is never crossed with canonical systems, it is no longer nearly specialised, has the flexibility and adaptability to various fields, and accidentally, it is a model of emergent phenomena of DNA usage in live cells. However, the generative proprieties of the L-system allow only one-way process. When the shape is generated, it is not possible to derive what L-system was used to generate it analytically. This applies to the genetics, as there are a few clearly expressed genes that their codons are known, for example, the colour of eyes or blood type: many proprieties, which are known to be inheritable, do not have the genes associated with them or have a complex combination of genes. Evidently, in the case of known fractals, the shape could be simply recognised, but for all possible shapes that are given to produce the L-system, which can generate it, it is an open problem. The information about the properties of certain L-system that generate certain shape could give insight about its genotype; nevertheless, as it is impossible to derive L-system form the shape, it cannot be used for measurement. Although the L-system does not seem to be useful directly to model an urban form or social capital, knowing that the fundamental understanding can create leaps in models universality, the search could be directed to the fundamental models of biological systems in hope to find the one that do.

1.1.4. My cell, my castle: simulations that react to the environment

Although the L-system has achieved a leap in universality, it did not become universal enough to be used as a measurement tool. In fact, the problem of getting the L-System form the shape is still unsolved by the contemporary science. L-System accidentally mimics the molecular assembly dynamics in nucleuses of cells; the fact that this process is inherited by the growth of multicellular microscopic and large scale organisms leads to the universality in growth simulation. However, it does not allow the simulation of interaction. If the shape grammar would generate a tree with the ingrown branches, L-System would allow it even if it would be impossible in reality. A model that could include interaction would enable even more emergent phenomena that is cellular automata, and the examples that extend the L-system in such way were discussed in the previous subchapter. It can be used in very broad fields: crystal growth, biological pattern, and fractal generation, social ghetto formation in urban environment simulation, city growth simulation, Voronoy cells generation, skeletons and neural avalanches simulation.

There are numerous different cellular automata with a wide array of properties; however, for simplicity, it is possible to describe one combination of properties as the simulation of cells in rectangular lattice, where the state of the cells change for every iteration of the simulation, according to the defined rules and states of the neighbouring cells. In an urban setting, the cells can represent the state of the land plot; therefore, the simulation of a city can be created.

Cellular automata was invented in an attempt to simulate crystal growth by Neumann (Von Neumann, Burks, 1966). At the same time, a similar approach was developed for the image processing to classify a part of shapes by their elongation, using a propagating wave of active cells to find the median axis that the researchers named skeleton. The distance from the edge to the axis is measured by the number of cellular automata iterations it took to reach it (Rosenfeld, Pfaltz, 1966). Although the author does not classify his model as cellular automaton, urban racial segregation model falls into the cellular automatons description (Schelling, 1969a). Notable cellular automata simulation is a shell pattern of gastropod mollusc 1d binary cellular automata rule 30 introduced by Wolfram (Wolfram, 1983). As this simulation is one dimensional, the pattern is obtained by recording its states throughout multiple time steps and plotting them next to each other. The propagating waves, the forest fires emerge in a probabilistic cellular automaton named forest-fire. A very simple model with three state cells exhibits self-organized criticality (Henley, 1993). Propagating waves and self-organized criticality will be important aspects of this research, because they relate to the comprehension of shapes and urban environment.

Cellular automata is used to model the city growth (Batty et al., 1999), simulate the propagating waves of neural activity avalanches (Beggs, Plenz, 2003). However, probably the most famous cellular automata is Conway's "game of life", where using simple three rules, it was possible to create plentiful of patterns of interesting behaviours, for instance, computers or Von Neumann machine (Von Neumann, Burks, 1966).

In order to understand how the cellular automata works better, one model will be analysed in more detail. In cellular automata, the simulated world is divided into discrete grid of cells, usually initiated with random values or one. In order to obtain the next state of the cell, a certain set of rules is applied to every cell. Game of life is totalistic, binary cellular automata on a square grid with Moore neighbourhood (Gardner, 1970). Totalistic means that the rules for each cell uses the sum of neighbourhood cell states, contrary to the elementary cellular automaton, where the set of rules lists all possible combinations of the neighbour's states. Binary means that the cell can have only two states, i.e., in "game of life", it is usually named dead or alive, but on and off are acceptable. Moore neighbourhood means that 8 neighbours are considered: four in straight directions and four in diagonal. Von Neumann machine means artificial life, machine that could replicate itself. Although Von Neumann had implemented it on his own cellular automata with rule set specially tuned for this purpose, the earlier implementation of the game of life rule set is interesting because this rule set is more simple and universal, but the fact is that it took forty years to find it (Wade, 2010). Controversially, Conway hated "game of life" most of his life (Black, 2014). The rules of "game of life" are incredibly simple: cell survives to the next time step if it has 2 or 3 neighbours alive, dead cell becomes alive if it has 3 neighbours, otherwise the cell dies. The fact that such simple rules generate such a wide variety of emergent results is keeping the popularity of the "game of life" alive. Despite the popularity and universality shown in a wide range of applications, cellular automata models do not come without restrictions. Other configurations of this cellular automaton have been tried; there can be 512 rule combinations, many of the most interesting have been named, but there is only the one that produces such a wide array of phenomena. Most of the others die off, reproduce to fill the whole simulation space or stabilise in several time steps (Heudin, 1996).

There have been many attempts to search for cellular automata rules that produce interesting dynamic patterns. Increasing the influence radius or number of states can enlarge the rule-space dramatically. Then, it becomes impossible for a person to test all the rules, and the automatic search systems lack the function of evaluation. One interesting attempt takes a genetic programming approach to generate new rules and nearest neighbour evaluation of weighted statistics of simulation to learn the direction of a searchable pattern from an interactive user input (Heaton, 2019). Although this simulation is able to achieve visually interesting results, the rules are too difficult to interpret.

Cellular automata can be criticised for being not continuous; therefore, not appropriate to simulate in the real world, which seems to be continuous; however, it is not, at least on very small scales. The smallest live organism is one cell. The smallest amount of matter with certain chemical properties is a molecule, which itself is composed of atoms that cannot be divided without changing their properties. This goes on until the limit of the discovered scale is reached. For example, the photons are unique form of the matter that cannot be stopped; they exist only in motion, a decaying electrical field generating magnetic field next to itself, and the decaying magnetic field generates a new electrical field next to itself again. Analogically, there are formations of cells in "game of life" that converge itself to the starting state, but in different place slightly moved (Sapin, 2010). In the terminology of cellular automata, there is a term photon that is used to describe the structure that moves itself with the speed of one cell per time-step. Another example is the chemical reactions that are not continuous; given two pure chemical reagents, there is usually one way they will react, if at all. The proprieties of atoms are represented on the periodic table of elements, as the name implies atom properties, namely, valence, which is the ability to form bonds with other atoms by sharing electron orbits that are periodic (Менделеев, 1869). Although it is not called periodic, the plot of elementary cellular automata exhibits periodic properties (Obando, 2015). The theory of quantum foam states that there is a very fine grid of space where quantum particles can only occupy a certain space in a very fine grid, measuring one Planck length (1.6 x 10-35 m) has been recently confirmed (Vasileiou et al., 2015). It is common to express phenomena in nature with mathematical equations, although equations allow predicting some claim that equations do not explain what is happening. As simple rules of cellular automata can explain it better, there were suggestions that all science should be re-evaluated in the scope of simple rules (Wolfram, 2002). Some sciences already rely on simple rules to describe the phenomena, like chemical reactions or interactions of quantum particles (Feynman, 1949). Despite the universality of cellular automata, there are obvious limitations in difficulty of creating such models and requirements of computational power to run them.

In order to make a cellular automata model, one has to have a hypothesis about the small scale mechanism that would result in emergent behaviour that is known and recognisable. It has to be created by using human insight and skills, detecting the desired phenomena as well mostly relying on the human interaction. The resulting model has to run many time steps for the emergent behaviour to be able to appear, and for every time step, for every cell, the calculations have to be made, which sums up to a lot of calculations. Further on, some models, especially those that use cell with continues values, are highly parametric. Multidimensional parametric space emergent results are not guaranteed to appear under all parametric combinations. Sometimes, the result that is interesting only appears in a very narrow parameter space. In order to find it, many simulations should be run with slightly adjusted parameters in the process named "model sensitivity analysis", which even more increases the requirements for the computational power. Those types of models are useful for helping to understand the phenomena under inspection, helping to form further hypothesis, and deriving mathematical formulas, describing the appearance of emergent phenomena. Such models can be useful to test certain hypothesis, but not that much in interpreting the real world data.

Gene, L-system as well as the theory of Leyton's process grammar are linear, but the virtual implementations of genetic algorithms do not have to be, because graph structure is a subject to evolution (Poli et al., 2008). There is already a proof that the informational chemistry that is coded in the four nitrogenous molecules are not unique by creating a new system with 8 nitrogenous molecules (Hoshika et al., 2019). Further artificial genetic algorithms can facilitate and evolve the nonlinear systems like graphs or matrices. Therefore, in the simulation or the modelling of the urban form as memory, not going the full path from the form to the skeleton and further to the linear storage of information, like process, the grammar could be acceptable in the terms of simplification. Further urban form would not be composed of countable parts; the measurements of length and angles would be real values, as it is a requirement of the genetic assembly and L-system; therefore, the converting of urban form to linear form of information inevitably would lose some precision. Even though it could be organised as parametric and parameters could be tuned to minimise the loss of precision, by analysing the urban form in graph representation, the precision is not lost. It has been already mentioned that neural simulations exhibit the same progressing wave formations as the skeleton analysis of the shape. It is possible that there are more clues that suggest that space comprehension is connected to the shape skeleton; by analysing more studies with animals, it would be possible to find it. Despite the fact that this subchapter does not show promise of cellular automata use to find the relationship between the urban space as memory and social capital, it could serve as an introduction to the cellular automata simulations, because there is a whole branch of cellular automata model dedicated to the urban simulations, which will be analysed next.

1.1.5. Cell and the city: urban simulations in the cellular automata

The emergent properties of many cellular automata inspired many researchers to adapt it to the urban simulation. As urban development is an emergent phenomenon of complex system of human society; the economic growth and adaptation of the natural environment, simulations as well follow the same structure and try to construct a complex system out of simple rules based on the real observed phenomena. As a cellular automaton is considered as a candidate model for this research, in this chapter, it will be mostly concentrated on the experimental research in the field of urban cellular automata. It is theorized that the computational power of computers that are used for the simulations was the limiting factor in the choice of model hyperparameters and as consequence influenced the significance of the results; therefore, extra effort will be made to analyse and compare the parameters of different simulations, and the correctness of methodology is used to simulate cell behaviour.

Urban simulations in cellular automata start from the idea presented by Tobler and proposes the use of cellular automata square lattice model with the citation from *Moby Dick*. Although the paper does not report on experimental data, it offers predictive discussion of possible different types of neighbourhood; nevertheless, it does not use the current terminology, proposes to use Markov chain for modelling change of land use in the cells (Tobler, 1975). Even though, there was not found any articles reporting on Tobler's experiments with cellular automata, this idea has been cited many times and inspired many researchers.

There were several arguments made about the cellular automata simulation for urban development; the earliest of them does not use the real world data; therefore, as this research is aimed at the use of the real world data, the following analysis of the cellular automata simulation for urban development literature will be focused only on the examples that use the real world data. The paper reports an experiment on the cellular automata model that produces structures of multifunctional cell types that visually resembles maps of the cities. Simulations were performed on 50 x 50 grid where one cell represents 500 x 500 m land plot. Although there is no reasoning behind choosing those parameters, it is suspected that it is motivated by technical limitations. The cells can have 4 states: vacant, housing, industrial and commercial. The models were initiated with 12 x 12 patch of vacant cells and then randomly changed to random developed type with probability of 0.3. The simulation is performed using random function, aggregated by the influence of neighbouring cells. The influence is calculated by using weight matrix that is consisting of a set of weights for every combination of the state transition. The sets themselves have weight for every distance up to 6 cells in horizontal and vertical as well as fractional distances, resulting from diagonal directions, totalling to 18 distances. Overall, the models had 486 parameters, which were tuned manually. The repetition of weights in the weight matrix that are presented in paper indicates the oversimplification of primary set goal, probably due to the struggle to tune such enormous number of parameters by hand. The researchers state that the influence formula is designed in such a way that it causes influence of randomness to diminish as simulation progresses due to the neighbouring cell, acquiring more values, but no test results were presented to support this hypothesis. In order to evaluate the simulation results, the fractal dimension of the city boundary was calculated and compared to the historical data that was acquired from the fractal dimension calculations of historical data of cities: Berlin, Atlanta, Cincinnati, Houston, Milwaukee, which were calculated at 80 x 80 grid, reasoning that they would not fit into the grid of their simulations. Grid size allows more points in the fractal dimension calculations, which results in the increment of accuracy. In the fractal dimension calculation discussions, the nonlinearly in cell count at different scales was interpreted as cities being multi fractal, i.e., having two fractal dimensions. However, it should be argued that the nonlinearity was caused by the simulated city borders, reaching the border of simulation and in the result of that becoming linear. Although researchers do ignore this limitation of simulation, naming it “boundary effect”, they do not test or rise the hypothesis about its influence on the fractal dimensionality calculations (White, Engelen, 1993). Overall, this model could be classified as reaching towards the real world simulations, because its mechanisms are based on the laws found in the real world, although this process is intuitive and natural, it is not the only one available; sometimes, the adapted simulation of other certain process can produce acceptable results with much similar mechanism, and in result, it can offer unique insights.

One of such models is based on the diffusion limited aggregation, where the particle is introduced into the system in some conventional distance from the external contour of the starting city. For each step of the simulations, the particle chooses a random direction and moves to the adjacent cell in that direction. If it steps too far from the city, it is deleted, and a new particle is introduced, if it moves next to the cell belonging to the city, it “sticks” and becomes part of the city. The only parameter that contributes to the ability to adjust the shape of the city under the simulation is the probability if the particle sticks. Therefore, if at first available opportunity, the particle does not stick, it can penetrate into the boundary of the city border and fill the gaps that would form if it “sticks” at the first opportunity. This simulation is just one step further from the randomly changing city edge cells to be developed, but newly introduced additional hyperparameters allow enough control to achieve satisfactory results (Michael Batty, Longley, 1994b). The results of this research can be interpreted in two ways: the knowledge of the mechanism governing the plot change is so scarce that random chance is the best approximation, or fractal dimension measurement is too simple to be used as a good evaluator of simulation results.

Simplifying the simulation engine limits, it results to a narrower scope and allows testing only the selected hypothesis. More interesting are the simulations that start from the hypothesis that tries to implement known, observable or reported by the other researchers’ phenomena in the micro scale and tries to see if it produces comparable results in the macro scale. The behaviour of different social groups is a common knowledge, but the question is how it would influence the development of the city. In the essence of this urban simulation, there is a lover level social simulation, which simulates a population of two different types, i.e., white and blue collar. The simulated agents have unique wealth and use it to compete to use the same space with each other to satisfy the demand of accommodation. The value of the plot is determined by the voting system, in which the agents participate based on their preferences and land plot properties. There are small cluster cities simulated in one simulation, connected with roads, which play an essential role in the city structure. The cells in the simulation have multiple properties: empty plots, buildings present, owner’s membership to certain population that is based on the collar colour, availability of the base and service activities, access to roads. The agents are not bound to the previous locations and can relocate freely based on their financial power preferences, and the price of the plot, which are calculated considering the proximity of services, jobs and roads. The city expansion and shrinkage have elaborate scenarios, which happen with probabilities, if certain conditions are met. The simulation is mostly described in the mathematical equations with few tuneable parameters: travel distance, services–population ratio, population by collar colour ratio, urbanised land limit. The model sensitivity was tested by adjusting tuneable parameters to the defined extremes, while other parameters were locked, resulting in two experiments per parameter. The visual result of experiments looked similar and could be described as the star-shaped city with a little variation caused by the influence of various parameters. The shape of the city in the simulation is mostly determined by the starting point of which is always the same cluster of interconnected cities. The star shape city pattern is a well-known urban form caused by the preference of urbanising plots that are connected to better roads. Starting the city simulation from the same point increases the consistency and enables easier interpretation of results, but mostly, it is useful to demonstrate emergence. The question is how wildly results can change starting from the same point with different rules or hyperparameters of the model. The model results are too similar, and authors themselves acknowledge that the model did not produce unexpected results (Semboloni, 1996). The most of the potential of the model was ignored because when you have a simulation, it is interesting to perform model sensitivity analysis and see where in the hyperparameter space, the disasters are simulated or the simulation breaks.

Urban cellular automata model not always has to be predictive or use real world data; it could be a theoretical demonstration of observed phenomena, emerging from the known principles. Webster and Wu tried to emerge results of the simulated economical welfare seeking behaviour. In this model, the cell can be in two states: urbanised as residential or industrial, although the authors state that the number of states can be expanded. The simulation starts with randomly assigned states to all cells, and during every simulation, the iteration probability of conversion to every type of every cell is calculated, taking into account the density and type of neighbourhood cells; then, they are tested against random function to determine if the conversion was performed. The virtual simulation space consists of a square area of 100 x 100 cells, although the cell numbers were not revealed by the authors and had to be counted manually from the illustrations. The simulation was executed for 50 iterations, during which the numbers of different cells, numbers of converted cells and other variables that were used in calculation of probability of conversion were recorded. Some of the recorded variables, for example, conversion rate, which became stable, could indicate that the simulation reached equilibrium or converged as interceded by the authors, but the conversion rate did not reach zero and other variables, for example, the potential efficiency and cell profit were still noticeably increasing; therefore, the chosen number of iterations was unexplained (Webster, Wu, 1999). Variable meanings and how they are calculated will not be explained, because the simulation produces what can be achieved with simpler methods; therefore, a simpler model is better, namely the Schelling racial segregation model (Schelling, 1969b). Although Schelling only made one dimensional simulation, it has been repeated many times on two dimensions and produced similar results to the results of Webster and Wu. Moreover, there is no explanation what amount of time one interaction represents or what area of land cell size represents, although for this simulation, it does not exhibit great importance, because it only seeks to prove rather obvious conjecture that policies, which determine cell conversion, are resulting in different patterns of land use. In this simulation, the most important factor is the industrial type cell influence to conversion probability with symbolic pollution factor, which is applied to 8 adjacent cells, in CA community known as Moore neighbourhood, therefore, resulting in simulation converges to industrial and residential agglomerations in the pattern that is visually similar to the Shelling racial segregation model. However, the authors did not acknowledge this similarity or cite Shieling's work, probably because Shelling's works is in social sciences, and this is Urban simulation. Despite the shortcoming of this research, it is valuable, as it shows that even with the implementation of very elaborate economic systems, which are stimulated accurately with extensive and heavy calculations, the result is still similar to the one which is achievable by very simple model; therefore, it is safe to simplify it in order to be more understandable. There is an important lesson not to complicate the cellular automata models by trying to implement every known theory, because at the end, it could converge to something simple that could be done with simpler model. In this way, a complicated model that produces the same result only proves that the complexity was not essential. The ability to simplify model to achieve similar results could be the most valuable result of this research, even though it was not addressed by the authors.

In order to address the discrepancy of real world irregularities and commonly used rectangular lattice grid for cellular automaton simulations, the transformation of the real GIS data was performed. Naturally, one would make such simulation from the land plots, but in this case as an object, and the essence of simulation is building rent; building data was used as well. Using real world data, there is an additional step before the simulation can be executed: one needs to generate the graph; therefore, the authors chose to generate one with Delaunay triangulation algorithm. The simulation was based on running landlord and tenant actions in every cell. The simulation rules were based on Rent gap theory. Rent gap is a potential but usually not utilised worth due to the lack of maintenance and renovation during the aging and affecting the value of property. The simulation was executed for 720 iterations, one iteration representing one month, thus in total, 60 years (O'Sullivan, 2002). The authors state that the dataset had 514 nodes, but did not provide the number of edges in the graph, therefore, tracing the illustration provided in the paper and performing Delaunay triangulation in the vector editing computer graphics software package "InkScape" that produced graph with 562 nodes and 1669 edges (Inkscape, 2020). For comparison, 50 x 50 squares of the lattice grid with Moore neighbourhood used by (White, Engelen, 1993) would produce 2500 nodes and 7000 edges. In the regular lattice cellular automata, this step where one chooses the type of neighbourhood, the most common are von Neuman (including, East, West, North, South neighbours) or extension of it, i.e., Moore neighbourhood (extending von Neumann neighbourhood with diagonals SE, SW, NW, and NE). However, in urban simulations, it is common to use radial neighbourhood where the influence is calculated in several cells, according to the set radius, and the influence is as well weighted, according to the distance to the cell. In this research, the building centre points are connected by using Delaunay triangulation algorithm, and the generated graph is used for the cellular automaton simulation. The authors stress multiple times that the location is of most importance, and it determines the theoretical side of their research, whereas the choice of Delaunay triangulation algorithm for graph generation interferes with that. For example, if one would take 4 adjacent cells, forming a rectangle with Moore neighbourhood; then, the fully connected graph would be used in the cellular automata simulation, but in the same situation, if the cell centres were used to generate graph with Delaunay triangulation algorithm, one diagonal connection would be omitted, because it attempts to cover the area with triangles with as blunt angles as possible. One could argue that Delaunay the triangulation algorithm only emits the longer intersecting connections; nevertheless, it could be argued that the intersecting connections should be allowed, and the rules for making or omitting edges of the graph should be motivated by relevant research and not just by the availability of algorithms. Moreover, it is not known how the irregularity of the node edges count influences the emergent effects in the graph. In this research, the connections per graph vary between 3 and 12. In is not known which approach is better to use, i.e., regular or irregular lattice, and this work does not attempt to answer this question. It just shows that it is possible, and this is a very important and relevant step for the future research. Additionally, this simulation does not allow the urban structure to change, grow or shrink, because it only has the existing buildings, and the simulation changes their proprieties. It would be very difficult to generate building shapes with predictive power; therefore, using building shapes for cellular automata graph could be attractive due to the availability of data, but it limits the research to narrower topics, for example, selecting a part of the city that is protected by the heritage laws and assuming it does not change.

Another aspect that is not changing very much is the census territories; therefore, the choice of a big enough territory could come up with a significant number of cells. In order to calculate the neighbourhoods, the authors used a parametric model with changeable radius, which produced a graph by using centre points of census territories, forming radial neighbourhoods. The model allowed the changes of use, density of urbanisation, demand of zones and accessibility. Two historical maps masterplans were used to populate the data of the irregular cellular automata, consisting of 1015 cells (for fitness function, when calibrating model parameters foe). The model was run once for one step for the calibration purpose and another for the prediction of the future state, which accuracy could not be measured. The model had 48 parameters and was tuned by the particle swarm algorithm; it is a variant of random search. The learning converged after 70 iterations when the improvement of the model was less than 0.1%. The Cohen's kappa coefficient was used as the goodness of fit measure, reaching 0.7 pseudo probability of agreement of land use classes between the prediction of the model and historical data (Pinto, Antunes, 2010).

Although the artificial neural networks were invented in 1960, its usage started to progressively increase in 2008 (Hao, 2019). It was partly caused by the availability of relatively affordable computer hardware, which caused multidisciplinary research of various fields that were including artificial neural networks as an integral part of the studies. Considering that urban cellular automata were not excluded from this trend, it is imperative to analyse such example. In this simulation, the known inconsistency of land conversion, which is hard to predict, has a helper layer that is provided with the help of artificial neural networks. The authors used ANN to compute the attractiveness maps, which affect the cellular automata probabilities to change the land use from biophysical and infrastructure variables. Biophysical variables in this case are distances from the current cell to the closest every other type of cells that can be not urbanised, residential, commercial, industrial, institutional, services, recreational and water. Although there are 8 types of land in the simulation, in contrary to the artificial neural network in described architecture, the only outputs are only one probability; therefore, the authors trained 4 different models, which generated the probabilities of 4 transitions, which are always from not urbanised to certain types of urbanised. According to the probabilities of attractiveness map, two functions are used to compute the cellular automata state in the next turn; first, one tries to expand the existing monolithic function territories by using Moore neighbourhood, the other tries to start new territories in undeveloped land, which use a stochastic variable to finalise if the transition occurs. The artificial neural network was trained on 12000 samples, which were randomly selected from the historical maps of two dates: 1985 and 1999; the same strategy was used to form the validation set. Artificial neural network architecture for every transition is different only in the input neuron count, but there is no explanation why the authors decided that for the transitions, certain type, but not all information, is necessary. Artificial neural networks were trained for 20–30 epochs, depending on the type of land transition, but the error plot during the learning shows that the learning stopped at approximately 5th epoch; no explanation was provided about why it was continued to train after the error stopped decreasing. The mean square error was used to track the learning of models, but the authors do not provide a comparison of the different models accuracy or the analysis of training results; moreover, the mean square error is very hard to interpret, because it is dependent on the sample size and the normalization of variables. Considering that the information about the normalisation is not provided, it is impossible to interpret the accuracy of the model predictions. Only a useful fact from the training data is a fact that the validation set mean square error decreases together with the training set. There was only enough data, or more specifically, not enough coverage for the conversion to the residential type of land; the data with other transitions was scarce. Overall, the model size was 334 x 360 in cells, one cell measuring 50 x 50m square of land, covering the area of 11.2 x 18 km. Cellular automata were simulated for 10 iterations, one iteration representing 1.4 years, in total, 14 years. The rules of change were designed in such way that the model does not allow the city to shrink or repurpose land into different use. The cellular automata model was evaluated with fifth fuzzy function, allowing for adjacent errors (Almeida et al., 2008). Although the usage of fuzzy function is well motivated, the model should have been measured with Cohen's kappa, as it has been used for such models in the past. This way, the models could be comparable. However, now, it looks that the fuzzy evaluation was used only to increase the accuracy. The simulation result analysis lacks simple statistical comparisons between the historical data and simulation: the numbers of changed cells overall and in the separate land, use categories, the accuracy of every land uses categories. Although this research has limitations, they are mostly caused by the limitations of the time. The authors used mostly homebrew tools and adapted them to serve the purpose of their idea; this is a very difficult task to achieve. Artificial neural networks were used incorrectly: they are diagnostic measurements and the difficulty of the problems suggests that a simpler model would have performed better. In several previous example algorithms, governing the change of the cell states, had many tuneable parameters; therefore, this research could mostly be an inspiration to invent a new kind of artificial neural network architecture that is specialised to optimise the parameters of the cellular automata model.

Most of the analysed urban simulations concentrate on the growth of the cities because it is easy to evaluate manually if the shape is recognisable as a city, or the accuracy of the result of the simulation can be measured by using statistical methods. Most of the models of the cell mechanics use economic models, with one exception, where social model was used. Despite the differences, all of them were using the known laws of economic and social systems and testing if the iterative use of them will generate emergent results. Therefore, this approach could be used as a continuation of this research, but it seems that it is not useful for finding the relationships between the urban space and social capital. This type of model requires elevated requirements for the data that are not related to the urban form and does not provide measurements of urban form. However, neural net or other model that can reproduce growth of the city could be treated as captured genotype of urban structure, if having numerous genotypes of urban structures, one could juxtapose them to other variables, such as social capital. This is a valid and very interesting research question but different to the one in this research. Although it was not possible to find a cellular automata model that measures the urban form, there is definitely a possibility to create one; this possibility will be investigated in the continuation of this research.

1.2. The maze of mind: how the cognition of space is used by the living organisms

Cognition could be viewed as an emergent result of a complex system of the brain, trying to make a useful tool that uses sensor information to make predictions of actions that should be taken and possible outcomes of those actions. In the process of cognition, the hypothetical meaning is associated with sensor input, and together they are stored in memory; therefore, any information about the cognition of immediate space that can be perceived could be useful in the process of methodology formation, especially if it relates to the mathematical or statistical analysis. Analysing the studies of live organisms, the reactions to environment, the navigation is done in hope to find an insight into the best models of urban shape as memory. Until now, the literature review was looked at as a theoretical science and how the theories were developed an adapted throughout the history. In this chapter, the attention will be concentrated on the studies of live organisms, because it is easier to conduct the controlled experiments, and the experiment could not be done on humans due to the moral values.

Another form of life, neither animal nor plant, in fact, it has its own category, the slime molds were used to construct logic gates; furthermore, they were used to control simple robots (Tsuda et al., 2009). Slime molds are very interesting organisms, particularly the category that is called plasmodial by the virtue of being one organism, which can grow fast, up to 2 millimetres per hour and up to very large sizes reaching to two meters. They are the largest known one cell organism. They have two growth cycles: in first, they spread outwards in a propagating wave of the thin layer. If such propagating wave encounters another food source, it enters the second stage and forms a pipe between one food source and the other. This propagating wave is similar to the one seen in neuron avalanches, and the same strategy is used in Dijkstra's path when trying to find an algorithm (Friedman et al., 2012; Dijkstra, 1959). In one research, this propriety of network forming was exploited to solve the path by finding problems in mazes, by adding one food source and slime mold at the start, and the other food source at the end (Nakagaki et al., 2000). The pipes were formed in between the walls of the maze, which slightly resembles median axis. The propagating wave of some species of slime molds grows according to the power law distribution (Schwalb, Roth, 1970). In another experiment by the same researcher, the slime molds could optimize metro network of scaled down Tokyo, where the pipes were formed between the food sources, representing metro stops, resulted in a network similar to the actual Tokyo metro network that was designed be the engineers. Similarly how the Dijkstra's algorithm remembers the visited nodes, the slime mold leaves residue, which is used to mark the visited places without food (Reid et al., 2012). Phenomenally, it can overcome this residue if the propagating wave is guided to a trap. Slime mold cycles of life can be simulated with cellular automata by using the reaction diffusion formulas; in fact, some species of slime mold grow in patterns, resembling the reaction diffusion patterns (Siegert, 1991). Slime molds exhibit many concepts that are used in this research, being very primitive and probably primordial organisms, they point out to the idea that those properties are fundamental to life. There is a slight possibility that this organism is the visualisation of the most simple space comprehension mechanism that could be inherited by the larger forms of life.

Choosing a path in the crossroad is a fundamental step in the urban navigation; therefore, in order to be able to predict how the crowds would choose their path would be valuable. Some of the most predictable are soldier crabs (lat. Mictyris longicarpus). It will be used as the first example, because it has two elements that are important to this research: surrounding space, influencing the behaviour of groups of animals. They live in the areas where foods are common, and their life is dependable on the flood cycle. In certain phase of flood cycles, they gather in formations and crawl around the beach. The movement of formations is very dependent on the physical environment. This was noticed by Gunji and was investigated further. The researchers made little parts of mazes and tested how the angles and the number of connections affect the movement of the crab formations. Moreover, what happens if two crab formations arrive at the corridor junction at the same time. The researchers were able to design such mazes, which acted as logic gates (Gunji et al., 2012). Logic gates are devices that execute Boolean operations, replicating the behaviour of logic gates, which is often used as a standard of highly predictable deterministic systems. They are higher level abstractions of simple electronic devices, like transistors, and can be used to design the processors. Usually, this applies in electronics, but it can be made in optics, fluids, mechanical, with billiard balls, microbiology and with DNR molecules, then it is called bioinformatics (Macdonald et al., 2009). Logic gates are a good benchmark of determinism or predictability, considering it is a building block of computers. Although to build logic gates with humans would not be possible or desirable, at least, it is possible in this research to try to find some evidence of behaviour predictability, starting from the evidence in the previous research.

Rats are social animals that share more genome with humans than crabs; therefore, it is often used in experiments that could give insights on human behaviour. In one study, the rats, which in the course of experiment are exposed to unfamiliar environments, stumble for a moment, and at the same time, the hippocampus activity is recorded, which otherwise is only present in certain conditions, for example, place cells (Wu, Foster, 2014). This activity is correlated with the previous recordings or the same subject during the time of navigation. A complimentary research showed that the damage in mice hippocampus cripples the ability to navigate in Morris water maze (Broadbent et al., 2006). This research reveals the structure of memory usage during the navigation process. There is some evidence that the power-laws, fractality and skeletonization are relevant in the spatial comprehension. For example, rat hippocampus has specialized cells, which fire, when the rat approaches the centre of the hex grid cell (Carpenter et al., 2015). Moreover, there are different cell types, which represent different scales of grids. This grid aligns itself to the environment as best as it is possible; therefore, if the room environment is skewed beyond the angles that are normal to the hex grid, the grid becomes distorted (Krupic et al., 2018).

Although the hex cells are not detected in the human hippocampus due to the technology limitations, it was possible to detect other more advanced features in the brain areas that are responsible for higher functions. There is a diagnostic procedure for extremely hard cases of epilepsy, where sensors that detect the origins of seizures are placed directly on the brain (Tiwari et al., 2019). This technology could allow looking for evidence of environmental comprehension. Although it has not been done with humans, a study macaque inferior temporal cortex, part of the primate brain, crucial for the final stage of visual object recognition, has shown that

28% more neurons are dedicated to the skeleton representations, compared to the surface representations (Hung et al., 2012).

It is worth to notice that soldier crabs do not crawl in formations all the time. In other phases of their life, they act completely differently, for example, when they dig their burrows, they perform this alone and are competitive and territorial. One could state that when they act alone, they act according to the probabilistic theory, and when they form groups, they act according to the deterministic theory. This is reflected in physics: Newtonian laws are deterministic, but moving to the smaller scale down to the quantum level, the laws become probabilistic. This probabilistic theory was probably inspired by the work of Lawton and Nahemow. They produced a cross plot where one axis is competence and the other is challengeability of the environment. In this plot, there are areas corresponding to the individual competence and challenging the environments, where individual is comfortable or stimulated and able or unable to adjust its behaviour to the environment (Lawton, Nahemow, 1973). Even though the challengeability of environment is dependent on the individual's competence, maybe, it is possible to measure it theoretically by parametrizing the shapes. The difference of the challenge to orient oneself in the physical environment of the city should as well be detectable in the urban form. Therefore, when analysing urban form from the challengeability perspective, one could gain important insights on urban form as the memory concept.

1.2.1. Putting a number on cognition: evaluating the influence of visually perceived surroundings

Given the piece of art, shape, or environment, is it possible to predict how the humans will comprehend it. The methods that have been discussed above can provide many insights, offer ideas how to rationalise the thought process and can help to notice connections. However, they do not offer certain measurements that could help to compare the real world objects in a chosen perspective. In order to create a methodology to measure the urban shape as memory, the search should be started at the fundamental theories, describing the methods, related to the comprehension of surrounding perceived space. The aesthetic feeling is an important part emerging from this comprehension; therefore, the computational aesthetics could offer additional insight into the urban shape as memory.

The story of measurements in computational aesthetics starts with Birkhoff's aesthetic measure, where the aesthetic measure is calculated by dividing the order with complexity (Birkhoff, 1933). It is attempted to rationalise aesthetic; however, the order and complexity are the measurements of feeling; therefore, this formula only expresses the idea of emergent feeling of aesthetic, where the feeling of order increases it, and the feeling of complexity decreases. Moreover, this is only an opinion, because it is not backed up by any research involving human subjects or statistical analysis. Later, Brikhoff subdivides the order to vertical symmetry, equilibrium, rotational symmetry, relation to the horizontal-vertical network and satisfactory form. Some of those indicators could be evaluated without human error, but others cannot; further subdivision results in more biased components. In the presented examples with the aesthetic measure "evaluated“, without surprise, the most simple designs get the best score, confirming the idea of the author. Although this is an interesting attempt to rationalise aesthetics, and it is an inspiration for some more recent research (Akiba, 2013) where it does not reach the independence from the human bias.

Another research that is worth mentioning had made a tremendous effort to understand the proportions of live organisms, without aesthetics in mind, just pure mathematical understanding of shapes that living things become. Thompson employs statistical analysis and takes into account the visual parameters as well as weight and volume of the organism, discovers many statistical connections between the analysed variables and finds mathematical shape transformations, which describe the difference between the animal species and makes use of Fibonacci sequence and golden section proportions (Thompson, 1942). Although Thompson did not try to juxtapose his research to human cognition, aesthetics or use his finding in the design process, others did. Understanding is a secondary way of knowing, according to Dixon (Dixon, 1986); therefore, understanding mathematical equations, which describe the shapes of nature and art, can result in aesthetic feeling. Circular dependencies described by Thompson were used to design contemporary furniture in the algorithmic generative process (Kilian, 2006); his book has shown to influence Avant-Garde art theory (Kaniari, 2013), and the principles described in it were suggested to be used in interstellar communication (Lemarchand, Lomberg, 1996). Mandelbrot refers to this book as the preface to his book (Mandelbrot, 1982).

In psychology experiments, human recipients could be asked to look at the picture and later fill the questionnaire or answer questions in the interview, depending on the research. This way, the statistical data could be collected about the answers. Unless the shapes in the pictures were geometrical and very simple, it is difficult to have any statistical data. In order to find a solution, Blum imagined a situation when one starts an imaginary fire in the imaginary grass (Blum, 1964). As this imaginary world is simplified, the fire spreads in all directions at the same time and forms an expanding circle. If one sets two fires at once not that far away, the expanding circles meet at a certain boundary. At this point, Blum probably unintentionally described Voronoy diagram (Voronoi, 1908) implementation with the cellular automata, which was discovered much later (Adamatzky, 1996). It is relevant to point out that Adamatzky describes this algorithm as a propagating wave of exited cells, the same process that has already been encountered in the skeleton generation and neuron simulation. Blum continues by starting imaginary fires along the shape one would like to analyse, where the expanding fires form the already mentioned boundaries in the centre of the shape that they started from, where they form a graph like structure named "median axis". Blum described methodology that analyses the graph of this median axis to acquire information that could be analysed statistically. An interesting fact of this research is that at the time it was conducted, no algorithms were discovered to get the median axis; therefore, a special optical device was constructed that could defocus the image in a very specific way for the median axis to be able to see and measure. Although Median axis allows simplifying the shape into a specific graph that is fit for the analysis in a certain way, it is still a manual process; all statistical values had to be evaluated manually; therefore, the methodology included lots of space for the human error and bias. Although later algorithmic implementations were discovered, some statistical values were lost in the process, because it was impossible to implement. Blum's method allows seeing the structure of the shape, but the previously discussed theories with golden section and Fibonacci sequence were interested in the proportions of the shape.

In order to get back to measuring shapes with proportions, it would be possible to count the details of the different scales in the method that measures fractal dimension. Mandelbrot wanted to model nature especially to capture the roughness. In Mandelbrot's mind, fractal means rough; he invented this term, deriving it from the word fraction; therefore, the fractal dimension is the representation of non integer dimensional space. An imaginative experiment could be considered where one wants to be twice scaled up piece of the wire by joining two pieces, then the mass of those pieces would be doubled. The same for the sheet of metal, but one has to join four pieces; therefore, the mass is increased four times, continuing with the cube where in order to get the scaled up version, one has to join eight cubes, and the mass increases eight times. One could probably see that in order to get a scaled up version of the object, the mass increased to the power of dimensions of the object. If one wants to scale up Sierpinski triangle made of wire, one only needs three copies to make the scaled up version. However, there is no integer power of two with result of three; therefore, Sierpinski triangle is not an integer dimensional object. It is possible to calculate the exact dimension with logarithm two functions; therefore, $\log_2(3)=1.5849$. Although this non-integer dimension is not real, all calculations are imaginative; it is still a useful measure for the comparison of different shapes. The method presented above is only adequate to extend for the calculation of fractal dimensions for self-repeating shapes of independent scaling factor, but most objects in nature are not self-repeating, as for them, the calculation of fractal dimensions is possible with different methods, e.g., Hausdorff dimension, packing dimension, information dimension and the box counting dimension, to name a few. They all follow the same idea presented before, and the calculated parameters are roughly comparable; therefore, only the former will be discussed. In box counting method, the shape or more technically polygon is drawn on the grid, and the rectangles of the grid or boxes that are touching the polygon are counted. This is repeated many times with varied grid size, and the grid size scaling factor and count of boxes that are touching are collected to the data matrix. Both variables of the data matrix are logarithmically transformed, and the linear regression method is used to find the slope of a linear function of transformed variables, which is the fractal dimension, or more precisely, box counting dimension. Using such method, one is able to calculate the fractal dimension of any shape; therefore, another level of research is enabled to juxtapose this parameter to the other testable variables and find new connections.

The fractal dimension was as well used to compare the urbanised areas of cities in the historical development perspective. The measurements were used to fine-tune city development simulation based on the diffusion limited aggregation model that was implemented in the cellular automata.

Hagerhall used Electroencephalography machine to measure Alpha, Delta and Beta waves of the human brain while the recipients were looking at fractal images of varied fractal dimensions (Hagerhall et al., 2008). Alpha and Beta waves are associated with attention and focusing, Delta is associated with drowsiness and deep. The author claims that the results show fractal dimension 1.3 being "the most restorative and relaxing".

Another way to see fractal dimension is to view it as the representation of the complexity of the shape; this could be beneficial in the continuation of Brikhoff idea of aesthetic measure. The ability to abstract the environment could be challenging. Furthermore, the challengeability of the environment could be perceived as a complexity of environmental space. The fractal dimension could be one of the measures of complexity, although it is more often used to measure the complexity of the computer generated fractals; nevertheless, it is as well used to measure the complexity of coastlines, street networks and market dynamics (Mandelbrot, 1967). Since fractal analysis acquired many techniques, such as box-counting, ball packing, ball covering (Falconer, 2004). They all have the same essence. For example, the box overlay method divides the analysed shape into the cells and counts the number of the cells into the shape fragments. This is repeated several times with different cell sizes. The resulting data is used to calculate various coefficients that can be compared with each other and the reference coefficients. It is possible that this could be a model of complexity and challengeability of the environment. Fractal dimension has been shown to impact the mood of observers and has a deep root in the cultural history and natural sciences. The process of creating shapes is repeatable in simulations; however, there are applications that are more abstract, and fast methods are needed.

The fractals themselves are related to urbanism; several researchers have addressed to the city structure as fractal. The term fractal city was chosen to describe urban segregation, income inequality and other social problems of rapid urban development (Soja, 2000). Several aspects of the city: architecture, urban form, street network and even social networks, were discussed in the perspective of fractality using the correct definition of mathematical fractal (Salingaros, 2003). Its essence lies in two principals. Salingaros named it “universal distribution” and “universal scaling”. Universal scaling is the principal, which governs how building facades should be subdivided. The subdivisions should be in certain proportions where smaller elements should be repeated more than the big, proportional to the size. It is a simplified description of fractal, and the author does not hide it; in fact, the fractals are the main motivator for his reasoning. His idea could be simplified to the following: architecture should follow the same principals as the nature. This is even more empathised with non-mathematical or rules of thumb, which as well draw examples from nature and argue how humans are evolved to perceive such visual environment. Therefore, architecture and urban structures build in in this way should encourage the thriving of communities.

Fractal dimension has proven an important measurement tool, which can be used to evaluate the architecture and urbanism. It as well has a relation to the urban space as memory, as it harnesses the experience of prehistorical ancestors, which is encoded in the DNA as the essential memory how to distinguish safe and dangerous environment. Although it should be used as a method of this research, it alone is not enough, because there are many studies, which verified this theory; therefore, the search must continue for additional theories and methods that could be incorporated with this theory for a synergetic effect.

1.2.2. Aesthetic forecast: machine learning uses cases for the prediction of influence of visually perceived surroundings

With the rise of Artificial Intelligence (AI), or more correctly, machine learning methods that were enabled by the availability of computing power, several attempts were made to facilitate the AI for the aesthetic evaluation of images. The increasing popularity of internet enabled the formation of many online communities; some of them were dedicated to art, sharing pictures that were taken with the camera. Such websites could be a great source of data, because online community members are already interested in the subject, the system of exposition of images rating and commentary collection are established; the user identity is guaranteed by the account authentication, and there is no need for artificial insertion of questionnaire to gather user opinions. In one of such website, the users can upload pictures, add comments and evaluate by two criteria, i.e., aesthetics and originality where the dataset of roughly 3500 entries were acquired. The cross-plot of aesthetics and originality evaluations show a clearly visible trend, and the statistical analysis confirms the correlation of those evaluations; therefore, the researchers point out that the average user is not skilled enough or do not make the effort or not honest enough to differentiate the aesthetic and originality evaluations. It could be explained by saying that originality is a part of aesthetics; therefore, they are related. The researchers created several methods for feature extraction that could evaluate certain criteria by automatically analysing pictures in the dataset: exposure of Light and Colourfulness, Saturation and Hue, The Rule of Thirds, Familiarity Measure, Wavelet-based Texture, Size and Aspect Ratio, Region Composition, Low Depth of Field Indicators, Shape Convexity, that extracted 56 features in total, which were used to train Support Vector Machine to predict the class of the user's aesthetics evaluation. The model performance was evaluated on separate dataset, which was not used in the training, resulting in 62% accuracy (Datta et al., 2006). An interesting approach to extract the features is to predict zones, named salient locations of the picture, which attract more attention of the viewer (Lai-Kuan Wong, Kok-Lim Low, 2009). It is as well possible to predict the presence of a verbal description in the destruction provided by the user, who uploaded the picture. This is useful when modelling the data without the presence of aesthetic evaluation (Wu et al., 2010). Inspired by Birkhoff"s aesthetic measurement, the researchers have evolved a genetic algorithm to procedurally generate images whose Birkhoff"s aesthetic components were calculated automatically. The image complexity estimation was calculated with entropy over the (raw, including red, green, blue, and derived, i.e., luminosity) channels. The order was calculated as the inverse of time it takes for the fractal compression algorithm to compress the image (Li, Hu, 2010). Others have tried similar techniques and improved results by gathering more data, inventing more feature extraction methods and testing more Machine Learning models (Faria et al., 2013). With the discovery of the back propagation, deep artificial neural network (deep learning) was applied to model every imaginable problem, modelling aesthetics included. Deep learning requires huge datasets with handsets of thousands of entries; therefore, its popularity created a demand for various datasets. One of such datasets contained the categorized photos as well as aesthetic evaluations (Naila et al., 2017), which was used by the team of researchers to train the deep neural network with specialised architecture, to classify the photos by aesthetic value.

Although it is possible to train the machine learning models with the data entered by the users to a certain degree of accuracy, this is not the true measurement of aesthetics, but the model of public opinion of certain population could be useful to model the same population, but global usefulness remains unknown. The fractal dimension seems fundamentally strong, but the methodology does not produce strong results in the brain studies. Propagating wave that is occurring in Blum's median axis is as well present in the neural simulation and skeleton generation research. Some of the studies discussed in this chapter have better continuations than others; however, for this research, it is more important that the theory would be tested on human subjects.

1.2.3. Memory of urban space: connecting the research of other animals to support the evidence in human studies

It is often believed that the behaviour of humans and other animals are unpredictable; it has been argued that the social sciences are not really a science at all (Gareau, 1987). In this chapter, the examples will be reviewed that show otherwise. The predictability of live systems, and most importantly for this research, the social systems, is the fundamental theory that enables this research. There is a possibility to find the influence on social capital by the urban space as memory. In the previous chapter, there were reviewed the skeleton structure and propagating wave that is used to create it several times in psychology and neural science; therefore, this could be a candidate model for the urban space as memory. In order to evaluate its eligibility, it should be analysed how universal is this structure. Some studies in more fundamental subjects, in order to find evidence of visual environments influence to animal brain, should be reviewed. Such studies could offer insights on the question: how the urban shape as memory can influence the social capital, for example, the evidence of comprehension of space and how it is influenced by the surrounding environment.

Le Corbusier argued that human actions are simply provoked by the function of surrounding objects (Corbusier, 1923). Therefore, Lang would categorise this phenomenon as deterministic, when the outcome is only determined by the circumstances. This could be true in the case of the soldier crabs, where simple system was needed to evolve for every individual to be able to know how to act in every moment. Lang as well has an opposite theory, which states that it could be completely the opposite: the behaviour of an individual is not determined by the circumstances at all; therefore, it is random or only based on the free will. Although the theories are completely different, they are compatible; they come together in unified probabilistic theory, where the probability of individual reacting in some way is determined by the environment and personal attributes of the reacting individual (Lang, 1987).

The previously discussed fractal dimension has some studies with supporting evidence. The patients with the view to the trees had better evaluations from nurses, took fewer doses of pain medication, had lower scores for minor postsurgical complications and had to stay for a shorter period in hospital (Ulrich, 1984). Although this is not directly linked to the fractal dimension, there are studies where tree health is diagnosed with the measure of fractal dimension (Murray et al., 2018). A complementary study in Japan recorded relaxation and reduced blood pressure in recipients who liked to sit near the wood cladding wall, and no significant response to those who disliked it; on the contrary, there were negative effects to people who disliked steel wall and no positive effects to the recipients who liked it (Sakuragawa et al., 2005). Natural things have a tendency to exhibit fractal proprieties; therefore, by measuring the fractal dimension, it is possible to predict health and the absence of emotional stress, then it has a significant impact on the comprehension of space and, perhaps, social capital.

Hippocampus is as well crucial for the navigation of humans, and navigation is impossible without the comprehension of space; therefore, logically, it as well takes in the comprehension of physical surroundings, and in the urban environment, it could be partly responsible for the influence of urban space as memory. For example, taxi drivers grow hippocampus during their practice: the research showed that the size of driver's hippocampus was proportional to the years of practice (Maguire et al., 2000). There is a complementing research of a unique patient, a licensed London taxi driver, who had sustained bilateral hippocampal damage and lost his ability to navigate (Maguire et al., 2006). These studies confirm that there is a part of the brain where the comprehension of space is performed physically. It has already been found in rats and monkeys, as discussed before; therefore, humans could have place cells and use skeleton representations in the compression of space. This is as well confirmed in studies about the space in memory. When given a test task on the computer without immediate feedback, the recipients could only complete it by planning, observing strategies; statistically, it was deduced that breath first search is used (Huys et al., 2015). Path finding can only be done on graphs; therefore, the result of this research suggests that subconsciously, everything that requires planning is subconsciously reduced to a graph. Moreover, breath first search, when visualised, exhibits a propagating wave of the graph nodes in "to test" stack, and the propagating wave, as it was already discussed, is found in neuron behaviour and suggest to be used in the shape comprehension. It has already been discussed that animal's comprehension of space occurs in the part of the brain called hippocampus. Some evidence suggests that brains evolved the ability to perform breath first search in the representation of space, which is optimised to a mathematical graph. It seems that the graph theory is good to perform the path finding in GIS systems and space syntax analysis as well as in human comprehension of space. Recently, the grid cells that are contained in hippocampus were discovered to be used in the functional magnetic resonance scanner and virtual reality simulation glasses in human trials (Doeller et al., 2010).

In the study of urban walk paths, Wang asked recipients to rank panoramic photos that were taken on couple routes, according to the safety and crime probability (Wang, Taylor, 2006). The results have shown "fear spots", which span across different respondent evaluations.

Cultured neurons on silicon strip detectors the same type that is used in CERN to detect the atomic particles; therefore, it was possible to record neuron activities (Beggs, Plenz, 2003). The neurons become active for a short time; this activity is called fire. Neurons fire in certain wave like patterns that are called avalanches. Neural avalanches appear in different amounts of neurons fired and durations; moreover, these properties have a power law distribution; therefore, it is stated that the brain operates in the super critical state. Cellular automata simulations of neural avalanches visualize the propagating waves. Straight skeleton is as well explained by a propagating wave; therefore, it could be a single logical state of the multistage model of a shape comprehension; the other missing component is the ability to see the contours of the world. Although in most cases it is not possible: there are edge detecting algorithms that extract contours from the real world image (Loffler, 2008). Moreover, deep artificial neural networks, sometimes, referred as a partial model of the human brain, have special layers to pre-process the image date that are named convolutional layers (Lecun et al., 1998). The neurons in this layer get information about all pixels and pixel neighbours, one pixel and the neighbours at the time. Convolutional layers solve the problem of data engineering or pre-processing for the image recognition problems by learning so called filters. There were already many filters invented and used; however, the convolutional layers are interesting because they learn filters from the data. It is not surprising that some of those filters adapt to the detected edges; therefore, it is safe to state that they compute silhouettes.

As the size of neural avalanches rises, the length rises proportionally as well; this phenomena is called universal scaling (Friedman et al., 2012). Phenomena that are observed in neural avalanches, i.e., the power law distribution and universal scaling, correspond to Salingaros' terms universal distribution and universal scaling, but the difference is that Salingaros talks not about the neural avalanches, but requirements of building facades (Salingaros, West, 1999). Different scales in rat grid cells as well contribute to these field findings. The similarity of terms is not a coincidence; the phenomena described by the authors in different fields are similar. Salingaros states that the facades that do not have these phenomena are biophilic, cause anxiety and illness in humans. By juxtaposing Salingaros' theories and the findings of neuroscientists, it is possible to draw a conclusion that the brain has evolved to process information that is presented in a certain way. There are many power law distributions and universal scales in nature and fractal systems. It is interesting to point out how the distributions of the brain processes correspond to the distributions in nature. If the straight skeleton model is a good representation how the brain processes visual information, it could be made even better if the skeletons were analysed in the scope of power law distributions.

Even though human behaviour is probabilistic, many phenomena can be predicted more accurately. Francis Galton found that the guesses of random recipients about hard to measure object is the average of the accurate prediction (Galton, 1907). Clapping in the audience synchronises rhythms in a few seconds (Kruszelnicki, 2012). The same phenomena has been observed in insects, for example, fireflies and crickets (Hartbauer, Römer, 2016). Even mechanical systems are known to spontaneous rhythm synchronisation, for example, metronomes synchronise in the same way as people on the millennium bridge in London (Sanderson, 2008). Such oscillations are the proprieties of a complex system that has many interactive paths, and the exhibited collective behaviour can be simulated by the agent-based system (Newman, 2011). The most famous of such simulations are flocking birds (Lorek, White, 1993). The most common applications of such simulation include spontaneous traffic jams that area known as traffic shockwave (Blatnig, 2009). Pedestrian dynamics form distinct lanes, governed by the simple rules, i.e., match direction, match speed (Appert-Rolland et al., 2011). In essence of such simulations are simple rules agents that have to be followed to achieve emergent results. Sometimes, this result is the observed phenomena, but other times, it is not observed, but desired result. Such as football tactics of triangles, where the real human players follow simple rules to achieve the emergent result during the game (Moura et al., 2013). While analysing London riots in 2011, they were political, and after that, the copycat riots concentrated on robberies (Davies et al., 2013). They had pulsing manner, the relationship between the distances of travel to the riot had power-law distributions, which is a similar behaviour to the shopping tendencies. There has already been found some evidence that shape skeleton, produced by propagating wave, can be used as a model of cognition of the surrounding space; therefore, concentrating on skeleton applications in cognition studies should be the next logical step.

Analysing processes in the brain can seem far from the architecture or urbanism. However, there has been found the same processes and laws: propagating wave, power law distributions and skeleton representations of shapes; therefore, they are most probably related. Because of this connection, the skeleton representations of shapes become a strong candidate for the evolution of methodology for urban shape as memory research and must be analysed further.

1.2.4. Skeletons in the closet: median shape representations as a step in visual comprehension of the immediate space in human and other animal trials

As the shape skeleton is a candidate for the evaluation methodology of urban space as memory, the next logical step would be to review its origin and universality. Universality is very important, because there is no hard evidence that skeleton representations are used in the comprehension of urban environments, although there are plenty of studies that cover many neighbouring fields; therefore, it is most likely used for the comprehension of space as well.

The history of straight skeleton invention starts with the effort to overcome the inability to statistically compare shapes that are used in human and other animal trials (Blum, 1964). Although there were other models that were proposed by others, Blum argued that his model was more robust and based on the findings of recent trials of the time. He introduced wave propagation system, which, when it propagates from the polygon, intersects itself and produces, as he named it, medial axis. Blum's analogy of propagating circular fields was grass fire, and he gave an illustrational example of fires starting at two different points in the field and propagating as circles and at some point, meeting to form medial axis. In case this experiment should be repeated, not to say that it should be, even if it was possible to stage the grass fire in the windless weather where fires would spread evenly after the experiment, the medial axis would not be seen, only the scorched land. However, it is possible to make a simpler and morally acceptable experiment to produce this propagating pattern with resulting diagram that is seen in polarised light by growing L-Ascorbic acid (as well known as vitamin C) crystals from methanol solution (Ito et al., 2003). However, making this pattern from point sources produces Voronoy diagram, a related shape to skeletons; therefore, in order to produce skeletons propagating wave, it should be launched simultaneously from all polygon of the shape. Blum suggested using a special optical device to produce medial axis from shapes.

This method is not used in the urban form analysis; however, there are numerous indications that it is connected to the way humans perceive shapes and space; therefore, it could be better to model space that provokes social interaction, moreover, the social capital. It is the method that takes any straight line polygon and converts it into a graph, which then could be analysed with the graph theory methods. Polygons without islands, or holes, usually produce a tree type graph. It is a sparse graph with low connectivity, where nodes have only 1–3 arcs. Surprisingly, it exists naturally in various fields, and some tests suggest it subconsciously by humans to process the visual information.

It is supposed that the most fundamental use of straight skeleton is in the field of mathematics, called origami mathematics. It is used in fold-and-cut theorem, which states that it is possible to cut any straight-line polygon out of the sheet of paper with only one cut, if provided that it is folded in a certain way first. There is a method to calculate the folding lines for this theorem, and the first step is to calculate the straight skeleton of the shape, which produces half of the fold lines.

River erosion in the fields of uniform geological material produces canyon with equally angled slopes. When observed from the birds view or orthophotography, such canyons ridges in the slope would overlap with the straight skeleton drawing (Held, Palfrader, 2017). Further skeleton approximation is used to explain earth topographic structures (Jiang et al., 2018). Moreover, the multidisciplinary research combines the uses of shape grammar to describe the changes in the land erosion in topographical model that uses median axis theory and straight skeletons (Kim, Cova, 2007).

In ancient times, the stonecutters used to maintain a certain angle of the cut with the original surface to carve letters into the stone. This was more efficient to cut, structurally effective and easier to read than the varied angle cut. Again, the constant angle to the original surface produces ridges, which, when observed from a long distance or projected back on the original surface, produces straight skeleton drawing. It seems like an extension of the letter strikes ends, named serifs, that are extrapolations of the straight skeleton. One theory states that serifs originated in the process of correcting little errors of stone cutters craft at the ends of the letter strokes (Goldberg, Romano, 2000). Recent studies close the circle by investigating the readability of the digital typefaces using the shape skeletons, similar to the straight skeletons, but generated using the cellular automaton (Beier et al., 2018).

One of the simplest pitched roof designs is to make slopes down to every wall, and maintain the same slope angle of every roof plane. In that case, the ridges of the roof construction will form a drawing of the straight skeleton in the plan or a bird view (Held, Palfrader, 2017).

In Computer Numerical Controlled machining process, the optimal path of the milling tool is calculated by using straight skeleton mathematics (Huber, 2011).

Deep learning scientists made an effort to create the artificial neural network to include the representations of shape skeletons (Demir et al., 2019). This could be a promising improvement, because currently, the convolutional layers of deep learning network do not regress neurons that represent the shape skeletons (Yosinski et al., 2015). The skeleton system is used to parametrically generate fonts (Egli, Crossland, 2013).

Straight skeleton 3D implementation is used in Discrete Element Method to optimise the complex object shapes by describing them only with spheres placed on the straight skeleton, produced from the original object (Mede et al., 2018). There are even developed 3-dimensional skeleton generators (Cornea et al., 2005). The usage of 3-dimesional straight skeletons proposed the thickness measurement of 3d-scanned products (Petrović, 2010). Further human recipients could recognize animal species from only the 3-dimendional skeletons of their volumes (Marr, Nishihara, 1987).

There is enough of strong evidence to support claims that skeleton is a good representation of human cognition of visual space; however, there is not so much evidence to confirm that it is as well used in human cognition of surrounding space. Therefore, the review of seemingly not connected cases shows the shape skeletons universality and strengthens its case for the usage as human environmental cognition model.

Straight skeletons are common in many natural phenomena and used in many fields. There are several important cases that are discussed.

The ability to recognize the shape constructed from Gabor patches and displayed for a very short time increases, as the fixation point of that gaze approaches skeleton of the shape (Kovács et al., 1998). Gabor patch is considered a common building block when working with visual stimuli. It is usually used on grey background. It looks like small blurred striped circle, but technically, it is a sinusoid convolved with a Gaussian. Human recipients could recognize the differences in the skeleton structure better than in the contour (Lowet et al., 2018).

A tablet PC with polygon on the screen was shown for random people in the New York City's Times Square and asked: "Please tap the shape, anywhere you like". Interestingly, the tapings aligned with the straight skeleton of the shape, although the skeleton was not shown, and the tablet PC screen was wiped clean that the next recipient would not be able to see the fingerprints on the screen. As if the straight skeleton was computed subconsciously in the recipient's brain (Firestone, Scholl, 2014).

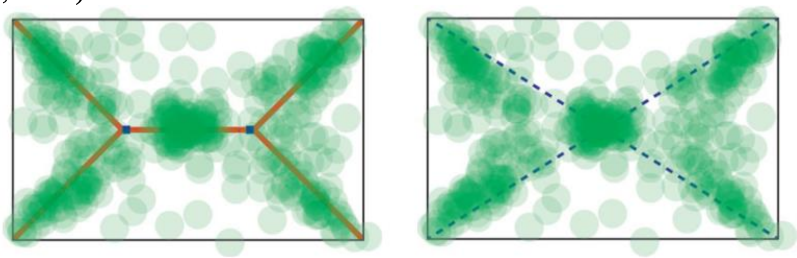


Figure 2. Shapes, medial axes and aggregated touches from experiments 1 through 3; solid red lines indicate medial axis, adopted from (Firestone, Scholl, 2014)

This is a simplified continuation of much more involved research where people have elaborated vision skills and can classify plant and animal shapes with high rate of success. Wilder have collected the sets of animal and plant shapes and calculated the skeletons for them. Data matrix, consisting of statistical parameters from the skeleton graph analysis and shape class, was assembled. Bayesian classifier was trained and validated for 80% success rate. For comparison, non-expert human recipients could classify for average 88% accuracy, the best being 95%.

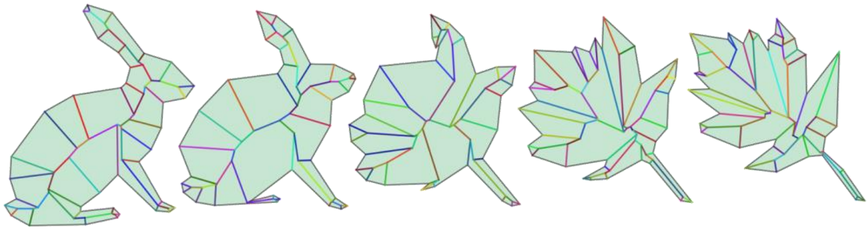


Figure 3. Polygon representations of animal (on far right), plant (on far left) and morphed interpolations (in between)

However, the real test was to see how humans and AI would classify the unknown shapes. The shapes of not animal and not plant were needed to get them to plant shapes where merged with animal shape, using the procedure often used in computer graphics, named morph. It parametrically produces one shape in between the two other. The parameter sets similarity to the either of them. New, morphed shapes were classified by human recipients and AI, resulting in statistically significant similarity (Wilder et al., 2011a).

Similar to the skeleton structure, "shock graph" was one of the most significant models, testing the correlation with neuron activity that was obtained with the functional magnetic resonance scanners (Leeds et al., 2013).

The skeleton is the continuation of Blum's median axis idea; further on, it is a fully automatic model, which does not require human expertise and does not suffer from human biases; therefore, it should be investigated further.

First time this shape was called skeleton by Rosenfield, who implemented Blum's idea on the IBM 7090/94 in Fortran (Rosenfeld, Pfaltz, 1966). They used image processing approach, calculating propagating wave steps on a square lattice, which could be classified as the cellular automaton, although the authors do not use this term. The emergence of propagating wave in algorithms of several related subjects, most notably neural simulation, has already been encountered. The abundance of this property adds to its significance as the core process of the brain function, including the human brain, as it is used to process the visual information, possesses the propriety of universality; therefore, it is most likely to be used for the navigation and comprehension of physical environment, including the urban space. This process of shape skeleton generation is based on the cellular automata; it could retain all features that were described by Blum. Rosenfield as well points out that skeleton possesses all information needed to reconstruct the original shape. However, the cellular automata models have weaknesses, because they have to simulate every time step of propagating wave and are only able to find rough approximations of the shape skeletons, as they are limited by the selected cell size and limitations of the chosen neighbourhood types. In order to overcome these weaknesses, there were developed other methods, laying their foundation in pure geometrical description of shapes and mathematical ability to predict the behaviour of a propagating wave by analysing and mathematically describing its processes.

The first results of such approach were obtained by Kirkpatrick who implemented vectorised (not based on cellular automata) skeleton calculation by using Voronoy diagrams (Kirkpatrick, 1979). This implementation loses the "speed of propagation" propriety that is described by Blum. Pure mathematical, used by the origami mathematics, "straight skeleton" (as well not based on the cellular automata) was described and implemented by Aichholzer (Aichholzer et al., 1996). This implementation, despite the "speed of propagation", as well loses the curve support propriety, described by Blum, hence, the name "straight". This "straight skeleton" method is used in CGAL implementation and in this research (Cacciola, 2004). However, the evolution of skeleton does not stop there, in order to regain some lost proprieties described by Blum, not parametric skeletons were upgraded to have varied slope angles (Biedl et al., 2015).

Another more robust attempt to implement useful parameters is to define a probability that skeleton arc is a good representative of original shape's part (Feldman, Singh, 2006). They state that the shape is data; the skeleton is a model. The skeleton is actually the generator of the shape: by calculating skeleton of the shape, the generator is found, its slow take about ten seconds per shape. The skeleton being the simplest explanation of the shape ties this theory to the widely accepted Gestalt idea of Pragnaz goodness of form (Sternberg, 2003). Gestalt theory allows the process of perception to be deconstructed into its elements, like acquiring a skeleton from the shape. Being a probabilistic model, Bayesians skeletons can be used to reconstruct the obscured shapes (Froyen et al., 2015).

An addition to the cellular automata type skeleton generation is a shock graph. This follows the same ideas as Bayesian skeletons, which traced contours are too detailed and produce too many small branches, touching vertexes in the traced polygon. In case of cellular automata, the nodes of the produced skeleton graph are on the lattice grid, contrary to the arcs of the graph, in which, only the end and the beginning are on the lattice grid. As it is possible to reconstruct the original shape of the skeleton, it is possible to state that cellular automata type skeleton is vectorization.

Although Aichholzer's implementation is the most pure and fast, it is not the fastest. Cheng proposes to calculate "motorcycle graph" before straight skeleton, which would help to predict the split and edge events and speed up the straight skeleton generation (Cheng, Vigneron, 2002).

Meanwhile, there was some evidence in human trials that positions near the medial axis of the shape are easier to notice (Wang, Burbeck, 1998). Notwithstanding that they claim to reduce the complexity and increase the speed, others argue that Aichholzer's implementation was of the same complexity (Huber, Held, 2011).

The discussed cases, i.e., the ability to model recognition of silhouettes, unconscious generation, contribute to the idea for the straight skeleton to be used as human cognition of a space model. Although there is a hard evidence that propagating wave of exited neurons is a process in the brain that is used in the comprehension of surrounding space and consequently urban space as memory, its universality in neural processes and abundance in nature at several scales suggests that it is the best model. As it seems that the skeleton is an appropriate model for human cognition of space, it would be interesting to analyse other theory that as well uses “shape as memory” terminology, just without urban, in order to compare them and choose the best features for the model of this research.

1.3. Adaptation of Leyton’s shape as memory to the urban shape

In his book *Shape as Memory (The Information Technology Revolution in Architecture)*, Leyton presents a way of thinking about shapes (Leyton, 2006), which as well speaks about the human cognition of space and tries to model it. The theory of form as a memory states that all forms contain their own memory of origin and transformational history, which led to its final form. In the example, Leyton shows a circle and applies transformations to produce an amorphic blob. All needed information to achieve the transformation can be abstracted to symbols, similarly how the transformational matrices in linear algebra are represented by the letters. Then, the sequence of those transformational symbols, named "process-grammar", holds all information or instructions needed to reproduce the final shape that is starting from the circle. The more transformations there are, the larger is the memory cache. Leyton claims that maximising information in this way will lead to new and better architectural style. Leyton has developed several notations of his “process-grammar” to be able to produce different kind of shapes; however, the notations were all specialised for special cases, and there was no universal model. His theory is related to L-system and Blum’s median axis and consequential to the shape skeletons. In this chapter, there will be provided a very detailed analysis of the development of his theory in order to reject parts of it that have no foundation and are incomplete, therefore not suited for modelling. Further on, it will be shown that other parts are covered by the related theories, despite the fact that Leyton rejected them.

It is based on the research that Leyton performed in 1982, where subjects were asked to describe non-basic geometrical shapes. The research showed that most of the subjects try to explain the shape by describing the transformations one needs to perform on a basic starting shape to make the deformed final shape. For example, the description "this parallelogram is rotated, sheared, stretched square" follows the transformation sequence in reverse order, until it gets back to the basic shape, which can be described unmistakably with one word.

While trying to find details about the research that was performed in 1982, the bibliography of the oldest publication (that was possible to obtain) by Michael Leyton is reviewed (Leyton, 1984), in which several Michael Leyton’s publications (which were impossible to find) are self-cited:

Leyton, M.: A Unified Theory of Cognitive Reference Frames. Proceedings of the Fourth Annual Conference of the Cognitive Science Society, pp. 204–209. Ann Arbor, MI: University of Michigan 1982; Leyton, M.: Experiments in the Structure of Form (Research report). Department of Psychology, University of California, Berkeley 1983; Leyton, M.: A Theory of Information Structure. PhD Thesis. Department of Psychology, University of California, Berkeley 1984.

From the year when Leyton's PhD thesis (1984) is cited, it seems that the experiments (1982) were performed while Leyton was studying. Disappointedly, it was impossible to find it. Interestingly, Leyton's name is never decorated with PhD, even at his profile page in Rutgers University, it still remains unknown if he did fail to defend his thesis (Leyton, 2018).

In 1986, Leyton published two papers with similar names to each other and his PhD thesis (Leyton, 1986a; Leyton, 1986b). In one of them, there are some details about the experiments he performed. The last experiment in the publication should be reviewed:

Subjects were presented with the successive pairs of figures in Fig. 15-that is, (rotated parallelogram, parallelogram), (parallelogram, rectangle), (rectangle, square) and asked which member of each pair seemed to be "more unstable, transient, or likely to change." The conjecture was that the subjects would always choose the second member in each pair. This was corroborated with considerable statistical significance [n = 12]; all 12 gave the hypothesized order; therefore, significance was greater than the pairwise ordering on reference judgements; i.e., greater than $t(11) = 4.771$, $p < .0005$...

To summarise, twelve people were asked three questions: to point one ("less stable") of the two figures. The statistical research standards require at least 30 percipients, when research is very expensive or risky (Everitt, Howell, 2005). Although the answer to the question is oblivious, it is highly unlikely that all recipients would give the same answer.

Other variables should be reviewed as well, e.g., "t" means t-statistic, as well known as Students t-test. It used to test the hypothesis about the mean of a quantifiable variable of normal distribution, but the test was a pairwise comparison; thus, the statistic method is not used correctly. If it is assumed that the number of correct answers was counted per person; then, it is possible to use t-test. From the following statement: "12 gave the hypothesized order", it is possible to assume that all recipients gave the correct answers, and all received the score of 3. This value statistically is not analysable, because it has no variation; it is impossible to perform the t-test on it. The formula for t-test is the following:

$$t_{\hat{\beta}} = \frac{\hat{\beta} - \beta_0}{\sigma_{\hat{\beta}}};$$

where $\sigma_{\hat{\beta}}$ is the standard error, and if all answers are correct, the standard error is equal to zero; it is impossible to divide by zero. Nevertheless, Leyton managed to calculate the statistics.

T-test is purposed to use for disapproving the so-called null hypothesis about the mean of the observed variable, which is exactly the opposite of that, one really wants to prove. It is as well called conjecture; Leyton uses this term. His null hypothesis is that "subjects would always choose the second member", which is the wrong answer to the question. This is a correct hypothesis. In this hypothesis, the subjects answered zero questions correctly; this is as well a standard way to use t-test. If it is assumed that the subjects choose random answers (and can answer that the stability of figures is equal) with equal probability, in total, 3 possibilities per answer and three answers are possible, then the average score would be 1. If it is assumed that accidentally, someone got all the answers correct and someone all wrong. Then, the null hypothesis is disapproved with t=4.06, p=0.02. With this, the synthetic data where the average score per person is one, a similar value for t-test is obtained, as Leyton claimed to get with the average of three. However, the correct way to interpret these results would be that with 98% confidence choosing the random answer to the question, one would answer more questions right than 0. To conclude, Leyton made too small experiments to be able to use the data statistically, used incorrect statistical methods in incorrect way and incorrectly interpreted the results. It is safe to assume that his research is inaccurate and should not be relied on.

Despite uncertain results of the research, some ideas are interesting and useful. This research tries to look a little bit inside the human mind by revealing how it copes with irregularities of the shape in order to simplify it enough that it could be translated into the verbal explanation. Leyton takes this finding even further by defining the mathematical symbols that are representing transformations. For example, "M+"– protrusion, "m-" – indentation, "m+" – squashing, "M-" – resistance. Leyton states that the final shape of the object can be represented by the "process grammar", containing the elementary transformations as constituent parts. Then whole object could be described with one mathematical group, containing the starting object and all the transformation that are necessary to describe the final object. However, it is impossible to make rotated parallelogram (analysed in the first example) from square, using this set of transformations. The elements of "process grammar" that Leyton introduces could be (if it worked) mostly useful to describe abstract and curved shapes. Although several architectural styles are using curved lines for decoration, most of the architecture is composed of the right angles and straight lines, with few exceptions, i.e., secession and deconstructivism.

Based on this principal, Leyton analyses art where he reduces the painting to line art and only the analyses curvature of line shapes in terms of process grammar. Despite him not taking into account colour, texture and lighting, he loses a lot of information; just from the grammar process, Leyton can produce extensive analysis of curvature. Nevertheless, this analysis does not reveal any unexpected conclusions, which are obvious and clearly not related to the analysis of the curvature. Narrative meanings are assigned to the elements of the process grammar that are not consistent and mistaken with artists interpretations of human anatomy (Gray, 1858).

Leyton's path of translating his theory into the mathematical language could be motivated by the fact that the theory can be only reasoned, contrary to the theorem, which is a theory written in a mathematical language; however, the theorem can be proven as well. Nevertheless, Leyton does not provide any definite proofs.

In *Shape as Memory: Theory of Architecture*, Leyton suggests a style of architecture, which would maximise the amount of stored information by increasing the amount of transformations that are needed to acquire the final form (Leyton, 2006). Although there is no such architectural style as "destructivism" proposed examples, provided by the author, it could be identified as such style. In the proposed buildings, there are no indefinable functioning elements, like windows or doors; nevertheless, the author classifies them as administrative buildings.

One can look at the curvatures of the lines and mark the points in the curvature extremes with corresponding process grammar symbols. This is done in many illustrations of Leyton's publications. At the first glance, the process grammar does not ensure recoverability by failing to provide enough information to reconstruct a shape described by it. It does not answer the questions in what order transformations should be applied and where, and the process grammar allows little control of the amount of transformation, but is very general, and it does not carry enough information how to use it for the exact shape reconstruction. One can theorise that the capital letters representing the transformations are matrices that are containing answers to these questions, but this has still to be done because Leyton has never tested his theory or discuses about what it would actually take to make it work. A good way to test the theory is to build a simulation; therefore, Milios (Milios, 1989) cites *A Process-Grammar for Shape* (Leyton, 1988) as an inspiration for developing and testing working computer model to be able to reconstruct hand traced shape of the orthophotographed cloud. Milios experimentally finds that Leyton's theory produces unstable models and for his final model, uses other theories. Moreover, Milios dives deeper into the math, revealing the contents of transformations matrices, contrary to Leyton, who never discuses transformation matrix contents in his publications. Interestingly, Milios never tried to classify his model as universal; on the contrary, he pointed out that it has purpose to transform a circle into other abstract shapes. In fact, Leyton cites Milios article in the usage of his theory. Leyton's overgeneralisations of mathematics behind his ideas is as well noticed by (Hendrickx, Wagemans, 1999).

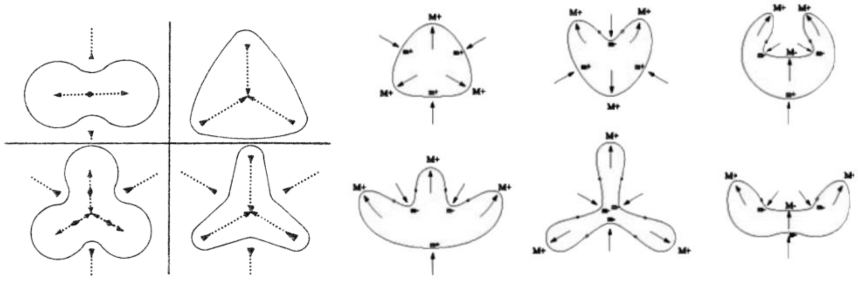


Figure 4. Comparison of Blum's descriptors of shape on the left and Leyton's process grammar on the right, adopted from (Blum, 1964) and (Leyton, 2006)

Leyton cites Blum, and incidentally, Blum's illustrations are similar to Leyton's, but the explanations are different (Blum, 1964). Further on, Leyton literally dismisses evidential similarities (Leyton, 2006).

In *A Process-Grammar for Shape* Leyton cites and uses Blum's idea of median axis, but in the later works, it is forgotten. It seems though that amorphic shape transformations lie on the axis, which they produce themselves. Although this does not explain what information lies within the Leyton's transformational matrices; however, it brings Leyton's theory closer to the straight skeleton, which evolved from the median axis analysis.

Leyton in his later publications provides multiple claims of his theory usage; it is cited in the experimental literature as a possible option to the model shapes, but not used (Feldman, Singh, 2006).

In one attempt to describe "classical architecture," Leyton abandons his process grammar for a completely different theory, which is stated as based on the process grammar, but it has no similarities. It is not based on the mathematical notation. It describes two dimensional shapes as a path travelled by a point and three-dimensional shapes as path travelled by two-dimensional shape. There is no reference to the software used to create illustrations, but investigating Leyton's personal site, some experimental illustrations were found, in which it is possible to recognize some distinct features of the software package "3D Studio MAX", namely, the default material colours and the special way it renders wireframes. Since 1996, "3D Studio MAX" has parametric deformations as well as a system named "Modifier Stack", which holds the history of every object’s transformations and allows user to change the parameters in every transformation interactively (3Ds Max History, 2011). It seems like Leyton is just poorly describing user interface of "3D Studio MAX" and claiming it as his philosophical proposition. In fact, "3D Studio MAX" was not the first software package to have this feature. Prisms had procedural object and material generation with graphical flow type language in 1986 (Prisms, 2012). It means procedural generation and animation of three dimensional objects happening in a product, when Leyton formulates his theory about two dimensional shape generation. First CAD (computer assisted design) machine named "Skechpad" that was made during the thesis of Ivan Edward Sutherland in 1963 (this is the time when computers were usually operated through teletype and punch cards) had unique and interesting parametric features, some of which are not present in the current CAD product. Namely, it could parametrically transfer the curved shapes into the straight line polygons. This is a working model, and Leyton is still struggling to put curved and straight-line shapes in the same theory in 2006.

Leyton’s idea to use an analytical model designed to describe shapes in order to model comprehension or the memory of the same shapes is not wrong and useful. In this research, the same approach of modelling is used; it is based largely on the same theoretical background. The slight difference is that they are made based on the better experimental evidence than in Leyton's approach and are more compatible with the statistical analysis. The slight changes that are made are not big enough to wander out of the scope of "shape as memory"; therefore, this term will be used.

In order to be able to compare urban space as memory to the social capital, it is necessary to have statistical measurements. It is already known that Leyton's shape grammar and L-Systems fall into the formal grammars category. In case there would be a method to determine the formal grammar that is needed to create the given shape, it would be possible to analyse the shape grammar expressions to gain statistical data, but such method has not been discovered yet. Moreover, converting the shapes to expressions of formal grammar is similar to achieving paramount compression; therefore, it would as well narrow the scientific question to the most challengeable urban form as memory. However, a more interesting question would be to openly explore all variations of environmental challengeability; therefore, the skeleton model, with its universality, similarity to Leyton's ideas, tested methodology, and available tools, is more suited for this research.

The skeleton representation of urban shape is most likely a useful model of abstraction, because there are a handful of models that work in architecture and urbanism; moreover, skeleton representations are used in the abstraction of visually perceived environment, which is directly related to architecture. Salingaros' theory of "universal distribution of scales" is applicable in both architectural and urban fields. Further on, in Alexander patterns, there are the same proportional rules applied in a wide array of different scales. Finally, human behaviour in certain situations in urban environment could be described in the same principles as the ones used for the skeleton representation of polygon shape, namely "propagating wave". Therefore, the skeleton representation is most likely a good abstraction of urban space as it is processed and stored in the human brain, and therefore could be named urban space as memory.

1.4. Ignorance of spatial influence

Up to this point, all references were not directly connected to the social capital or sociology, it was only assumed that they matter, but starting with this paragraph, the observations that are directly related to sociology will be reviewed. It has already been established that the human brain, as well as all brains, is a complex system; it exhibits non-linearity in the form of power law, criticality, spontaneous order in the form of a propagating wave of firing neurons; it adapts and has a feedback loop, which reacts to the environment. People, controlled by their brains, as well form complex systems, which exhibit criticality and spontaneous order; similarly, it was seen in London riots. However, not only negative, but social and economic organizations are in the same category as well. Moreover, human mega-creations, including power grid, transportation system and even cities, can be analysed as complex systems. Therefore, the complex systems in the structure of civilisation are influencing each other in various ways, resulting in the most optimal or healthy tendency, being inconsequential. The juxtapositions of knowledge and factors against the usage of knowledge should be discussed. Urban space is a part of the artificial environment, which has to be created by someone. In order to be a good creator, one has to study where to learn the most influential names of creators in the studied style or movement. Then, it is only natural that every creator as well wants to be remembered for some title. In order to create something new, it has become difficult and challenging, because all simple patterns have been enchased. For a creative author, this sometimes becomes a dilemma to repeat the well-known patterns that most viewers would understand and appreciate, and at the same time neglect one's creativity to explore new frontiers (Adorno, 1998). These save interpretations of the same patterns could be felt intuitively, and the medieval style consistency and revolutionary change could be explained by reusing the same appreciated style, until most patterns that could be still in the same style and original are exhausted, and the remainder of the patterns limits the creativity too much; therefore, a new style with new patterns is explored. On the contrary, it is difficult and risky to invent a new style; therefore, it is not done too often. The history of art spans many centuries; therefore, this exhaustion of patterns in the same style and style revolution happened multiple times up to the point where society was mature enough not stick to the same style for a whole century. Following this path, there are only that many variations of visual styles that do not make reptilian complex not to scream "danger" and enforce other parts of the brain to compensate and override. This compensation could be no that apparent, and it could be assigned to new labels that make it seem not that bad, and attention could be directed to the other emotions, like exiting or breath-taking. Moreover, it could be disappearing from the attention altogether, as people become accustomed to their environment. Cole notices that people do not see culture because they live in it, as a fish does not see water (Cole, 1996). Cole as well discusses Zeitgeist, spirit of time, the term introduced by Hegel (Hegel, 1807). Zeitgeist is as well used to describe the entirety of design rules that are used by historical architectural styles, therefore, the urban form as well (Saarinen et al., 2006). This is done in the scope of time and can be experienced only by studying architectural history, but in the everyday lives of ordinary people, the physical movement through a city or urban form is more common; therefore, the encountered spirits could be associated with the place. Although Cole does not discuss the spirit of place, such term is known and used. It originates from the ancient Rome, where it is named "Genius loci" and used to describe 265 unique districts, where the citizens had different deity, rituals and festivals (Woolf, 2012). The spirit of the place could have a positive influence on the people in it, like Alexander's social place pattern, but it as well could be negative (Alexander, 1977). As culture is not seen, people as well cannot realize how surrounding buildings affect them and influence local social capital. The cultural effects on social capital are as well noticed by Putnam (Putnam, 2001a). The cultural influences are not perceived, which change the probability to make new connections and grow social capital as well as maintain the existing connections and current social capital network. One does not realize the influence of the environment; therefore, such effect could be a context dependent psychological process (Cole, 1996). Social capital could be a part of another term used by Cole called national "Geist", the spirit of the people; therefore, in order to communicate the idea of this research, the urban form as a memory connection to the social capital, people only able to think only in mysticism terms, could be rephrased as the connection between the spirits of the people and the spirit of the place.

There are clear indications that the properties of urban environment influence the social behaviour of humans. The skeleton is not the only method that could capture the comprehension of urban space; there are methods that are built on the statement that social interaction is connected with the presence of people.

1.5. Meaningful dimensions of urban form: urban analysis methodologies based on the mathematical graph theory

The urban form and social capital have a common mathematical theory that offers methods suited for analysing both phenomena. It is a graph theory, which not surprisingly was invented by Leonhard Euler in 1736 to tackle urban problems (Euler, 1741). Graph theory deals with networks where nodes are connected with edges; it allows different types of edges and nodes and provides methodologies of graph manipulation. Euler's paper describes three situations how the cities are separated by a river, and islands can be connected back with bridges. In one example, the city of Konigsberg in Prussia was established in the location around Pregel river, covering two large islands. There were seven bridges, connecting different banks, land to the islands and between the islands. The problem Euler was trying to solve is to find a walk that would cross all the bridges only once. Although this problem had no solution, other imaginary cities were solvable. The analysis included a list of edges and nodes, contrary to the now widespread schematic graph notation. Currently, the graph theory is widespread and used in numerous fields, ranging from navigation, pathfinding, internet communication, analysis of infrastructure, social network and text; further on, in the field of artificial intelligence applications, there are specialized graph methods. Notable work, using the graph theory for transportation networks, made as a contract for the USA military, is presenting the basics of graph theory and introduces readers to graph evaluation techniques, matrix representations of graphs and even algorithms of graph generation (Garrison, Marble, 1962). Graph theory was used to simulate citizen migration in urban system (Riddell, Harvey, 1972). Graph models that are used to predict the urban traffic flow were evaluated with statistical parameters in (Ogunsanya, Ade, 1986).

As it was shown in the previous section, the social capital and urban space comparison could be true, because there are many references to the fact that the urban form may have a relationship with social capital. The further graph theory provides tools for urban analysis; therefore, it has to be analysed further in search of the suitable one for this research urban shape indicators. One of the most common ways of measuring urban form is the syntax of space, which was developed in response to the then prevailing grammatical theory of form (Hillier, 1989). Contrary to the form grammar, space syntax is not designed for the form generation, but is a diagnostic method for calculating a wide variety of indicators that were later linked to the street traffic (Hillier et al., 1987). Using machine learning, any function can be optimized; therefore, if one would treat space syntax as a measurement function, it could be used in optimization with another function that could generate street networks. There have already been made attempts by using Rhino Grasshopper extension with core to this process components "Galapagos", "depthmapX" and "SpiderWeb" (Schaffranek, 2013). Street polyline, aggregated by the angle of turn threshold, the lengths have been shown to have logarithmic distribution, which is as well known as power distribution and already have appeared in several papers, reviewed in this dissertation (Figueiredo, Amorim, 2005). The real estate sales and crime depend on various parameters of space syntax analysis (Hillier, Sahbaz, 2008) as well as the purpose of usage type of buildings (Zaleckis, 2018). These methods only analyse space as it is connected, accessed and visible. For example, spatial syntax axial analysis analyses the aspects of street axes or spatial axes.

In the real world, the street axes do not exist as well as graphs generated with shape skeletons; they are just a conventional term used for designing or representing a street network, map mapping, navigation or axial analysis. Although street axes usually influence the movement of people, they do not reflect squares, parks and other spaces where movement is free; however, the skeletons can be adapted as it will be shown later to analyse the shape of the buildings or the space between the buildings. Axial analysis is a good method for predicting the most commonly chosen streets during the human movement, but there is no indication that street axes should influence urban perception or social capital, although spatial syntax authors claim that mere human presence produces social effects. No one can argue that the way streets are connected affect the traffic in them; thus, these methods are that good at predicting traffic. Nevertheless, there are no indicators how the axial analysis is connected to the apprehension of the space or social capital.

The visual axis is generated from an empty space, not occupied by the buildings of walls. This method is often used to analyse the interiors of public buildings, although sometimes, it can be used to analyse the urban form, if it is constrained in some boundaries. This method has some limitations, considering the fact that it is not taking into account the 3-dimensionality of the space, a part of the distant object could be obscured and visible only from the top. The visual axis is generated from an empty space, not occupied by buildings of walls; therefore; it represents a line with the furthest visibility, ignoring how narrow is the field of view.

The same limitation applies to the isovist analysis, which analyses the 360 ° field of view, but it is almost never felt by a human. Meanwhile, the straight-form skeleton method has links to the form-grammatical theory as well, but it is diagnostic because a form-skeleton can be created for any shape; its graph contains enough information to reproduce its original form, and the form-skeleton itself is universal and often found to be fundamental. Although a very similar method is used for the urban visual axis analysis, this method is more suitable for single-storey interior space analysis, as it overlooks space and situations where a physically inaccessible object is visible due to its high height. This method is often used to analyse the interiors of public buildings, although sometimes, it can be used to analyse the urban form, if it is constrained by some boundaries. This method has some limitations, considering the fact that it is not taking into account the 3-dimensionality of the space, a part of the distant object could be obscured and visible only from the top. It has been shown that this is an important criteria in the study of perceived time passing (Yang et al., 2007). In this study, people were asked to walk through the predefined urban route with varied scenery without wearing a watch, looking at the clock or asking for time. At the end of the route, they were asked to tell the time, which had passed, only based on their feelings. Later, a virtual agent was put in the same route but in a virtual environment, and the distances to all visible objects were recorded in all directions. This is yet another method of space syntax family that is called isovist. It has been shown that using the data from 2-dimensional isovist was not possible to predict human subject answers, on the contrary to 3-dimensional, where answers could be predicted to a statistically significant degree. Although this study could point out that 3-dimensional isovist is at least partially modelling comprehension of space, it requires specific data, 3-dimensional model of the city. In some cities, this data is available, but not widely; therefore, this method is not available for this study.

Spatial cognition term as well appears in the space syntax research, although authors speak of cognitive space influence on the social indicators: their research only measures how well the participants are able to draw the maps from the memory. The authors as well used “skeleton” to describe the essence of their theory, although it is not the shape skeleton. Skeleton in this case is the preferred street, represented as its axis (Montello, 2007).

In some way, the shape skeleton analysis could be as well a part of space syntax methods, because skeletons can be generated from the data describing the urban shape; further on, it produces a graph representation of space; likewise, the space syntax uses a graph representation of space. It would be an additional tool with exceptional features: it is inspired by the process of neurology, and it allows the analysis of building shapes and shapes of the space between the buildings. The known space syntax methods have not been linked to the social capital or similar statistics; on the contrary, the urban thinkers have addressed the influence of urban space to the social behaviour several times; therefore, it is worth to investigate their findings.

1.6. Sociability engine: do social spaces generate social capital?

In architectural studies, there have been numerous attempts to capture the proprieties of urban and architectural spaces that encourage socialisation. Some examples should be reviewed to gain insights how urban space as memory connects to the social capital.

Supposedly, following the same idea, White in his studies of city plazas composed a list of plaza properties, which correlate with the popularity of the place. Some of them include: having places to sit, having places with shadows, even better if they as well have places to sit, having plants, even better if they are big enough to cast shadows where you can hide from the sun, having water of some kind, natural or a fountain, having places to buy food (Whyte, 2001).

One of these proprieties is confirmed by the other study in Baltimore, Maryland, where residents were interviewed to determine sociability and civil engagement of several neighbourhoods. Their multiple regression analysis model could explain 22% variation of the social capital (Holtan et al., 2015).

Nevertheless, the surroundings effect on the sociability of the place was noticed long before that. Alexander often uses the term "social space", which he uses to describe the environment that encourages social interactions (Alexander, 1977).

There are many statistical parameters to describe social environment, one of them that increasingly gained popularity recently is social capital. It can be personal or local. Personal social capital is the same as personal connections, especially those that could be utilised to gain benefits. This type of social capital could be criticised as a negative phenomenon, and in some cases, this is true, local social capital could be perceived as more positive. In order to understand what local social capital means better, a situation should be imagined where strangers move into a new neighbourhood. At first, no one knows or have friends. As the time passes, they could intentionally or accidentally meet, and possibly become friends. As this process continues, people acquire more quittances and friends or summarised connections. This could be as well viewed from the graph theory. Graph theory is a field of mathematics, often used to analyse social problems (Barnes, 1969). It is composed of node and arcs. There are accepted other names for nodes and arc; however, in this thesis, only those will be used for consistency. It is assumed that people are nodes and friendships are connections or arcs. At first, it starts with null graph, where there are no connections, and as the process continues, the connectivity of the graph increases, until it settles at some point or reaches full connectivity. A fully connected graph could represent an environment where everybody knows everyone. This is the maximum state of social capital, which can rarely be achieved in the real local environments.

In the course of the social capital history, several types of social capitals were defined. They should be reviewed in order to select the most suited for finding connection to urban space as memory. The first use of social capital is local to describe the propriety of the place. It comes from the times when education was not compulsory, Hanifan established a school in West Virginia, but the attendance was low. He made a survey, which results concluded to low trust. Therefore, he organised multiple events to gather people together: meetings, agricultural fair school, exhibit evening classes, opened library etc. As people became more acquainted, the school attendance increased. Hanifan named this process the accumulation of social capital (Hanifan, 1916).

Local social capital is as well used in this research. It is measured by the methodology of Putnam presented in the book *Bowling Alone* (Putnam, 2001b). This book became a bestseller, influenced many politicians and moved the measurement of social capital away from the academic circles to mainstream. It spread from the USA to Europe, and since 2002, Europe Social Survey is as well measuring the social capital of European countries with the help of local institutions by KTU (KTU Politikos ir viešojo administravimo institutas, 2017). Before this methodology was invented, the social capital was only measured by surveys. The survey is still used to measure it until now, for example, Europe Social Survey. As the survey is filled with questions that were given to ordinary people, its conclusions can be only derived from the perspective of the recipients. In Putnam's methodology, the social capital is measured by its tracks. Macro changes of the environment can be measured without including the recipient opinions or all recipients.

There are many philosophical and theoretical implications that urban form has a connection to the social capital; however, it is not measured statistically on the wide scale. Social capital is calculated with the method from the factor analysis, named principal component analysis. In order to be able to calculate it, four factors should be calculated: civic engagement, voter turnout, census response rate and non-profit organisations for every administrational unit. Civic engagement is calculated by adding religious organisations, civic and social associations, business associations, political organizations, professional organizations, labour organizations, bowling centres, fitness and recreational sports centres, golf courses and country clubs, sports teams and clubs and dividing them by 10 and by the population of 10000. Voter turnout is the fraction of the population that voted in the last election. Census response rate is the fraction of population that responded to the census questionnaire. Non-profit organisations represent the number of non-profit organisations divided them by the population of 10000.

As a limitation of the administration borders approach could be the analogy to Lynch's theory of image of the city where meaningful borders are found in the city, contrary to the administrational borders that often do not represent anything significant (Lynch, 1960). Although limitation still stands, considering that the available data is organised in administrational borders, the hotspots of social capital could form an irregular grid with Voronoy diagram; therefore, urban shape as memory analysis could predict social capital hotspots.

Some social capital theories take the available connections literally and analyse the social capital as a graph of the network of connections (Licamele, Getoor, 2006). They gathered 11,644 papers from 1959 to 2004, containing 11,554 unique author names and compiled a graph of all collaborations where papers were arcs and authors were nodes. Using the derived from the graph analysis parameters: personal and group social capital, benefits received and given, the authors could predict the future collaborations with 88 percent accuracy. When being the co-authors of a paper, it is easily traceable guaranteed connection, which is not that reproducible strategy of the research in other fields where the social capital is present. Although this research is interesting and provides useful model and insights, it is not universal.

Another research takes the same approach even further by including patent data and inventing algorithm, which analyses social capital graphs in order to predict the influencers and trends (Subbian et al., 2013).

Another research completely ignored the networking or graph modelling of the social capital. There were asked nearly 50 survey questions, some of which were open, to nearly 9000 recipients in order to determine the social capital strength and the community boundaries as well as the key personalities in them. Then, the interviews of 2822 key personalities were completed to determine the leadership level community boundaries (Sampson, Graif, 2009).

Another researcher points out that the social capital is only a subtype of unified capital. In his theory, there are three subtypes of capital: economic, cultural and social. The conversions are possible between the subtypes. Social capital has a propriety “volume”, which is measured in the size of mobilize-able network connections. The ability to accumulate and maintain social capital correlates with the size of the capital of all subtypes. He distinguishes economic capital as the root of other subtypes. The economic capital may be difficult to measure. He as well shows that the spare time is an indication of capital of all subtypes (Bourdieu, 1986).

The virtual networks are not always dissociated contrary to the popular belief. There are properties of connections between the members of a network. Individuals can gain benefits from the social capital, but they as well have to invest in order to be used by the other members of the network to maintain their connections. Some authors report negative effects, called “paradox of embeddedness” (Huysman, Wulf, 2004). There are even studies of the evolution of the definition of social capital (Adler, Kwon, 2002).

A similar approach was used to measure the social capital of the internet blogging sites (Smith et al., 2007). The researchers collected 13 million entries from 38 thousand blogs. There, the social capital is a hybrid graph with connections, representing different kinds of social capital, bridging and bonding, constructed by analysing the hyperlinks and compatibility of blog post subjects. Hyperlink analysis is straightforward: the connection strength is proportional to the times blogger hyperlinks another blogger’s post or mentions his name. The compatibility of the subjects is measured by corpus analysis using Latent Dirichlet Allocation (Blei et al., 2003).

The future progression of this type of research would be online social capital localisation. There was already some social research that does it, but without the social capital. One research used data from the media site *Twitter* and collected messages from New York City. The messages were analysed by using a classifier method from the natural language processing field. Using mood tags assigned by the classifier and temporal and spatial nature of the messages, the researchers were able to assemble temporal mood maps. Another research in Italy produced a "hate map" by performing the sentiment analysis on messages from the social media site *Twitter* (Musto et al., 2015).

For this research, the best-suited methodology has been developed by Putnam, it has a clear description and should be easy to replicate in any country where significant datasets are available. Certainly, it would have to be adjusted to the local culture, because not in every country people like bowling or golf. Besides, this method is very appealing, considering the availability of the data.

To conclude, urban space as memory is a model of comprehension of surroundings. Additionally, it is connected to the feeling of scale and proportions and topology of space and shape grammars. The feeling of scale was noticed and named by different terms: golden section, Fibonacci sequence, universal scaling, universal distribution and fractal dimension. Further on, it has been confirmed that the grid cells respond to a variety of different scales in mice hippocampus for benefit in navigation. The topology of space extracts graph of median axis, which has slight variations in methodology and naming, skeleton, straight skeleton, map skeleton. The skeleton implementation in cellular automata exhibits propagating waves, which has been found in neuroscience and sociology studies as well. Other implementations of skeleton are just faster approximations of the main model. Finally, the shape grammars are generative models of natural shapes, fractals and basis of proposed architectural style. From the scope of the shape grammars, it is possible to analyse the interpretation process of the DNA molecule in order to create protein molecules, but there are no studies that link cognition of shape or space to the space grammars. Considering this fact and the fact that there is no method for producing shape grammar with the given shape, this method is not considered.

2. MODELLING THE INFLUENCE OF URBAN SHAPE (AS MEMORY) ON SOCIAL CAPITAL

The ultimate goal of this research is to find and explain the connection between urban space as memory and social capital. The urban space as memory has several components; therefore, before starting to build the main model, the universality of them should be retested. Starting in the topology of space, which ultimately produces a skeleton of shape, the question is if it can model the city growth. The answer to this question would additionally prove its usefulness in modelling urban space as memory. Next, the universal scale or distribution of sizes must follow the power law distribution. It is possible to test it by building a model, which starts from the randomly distributed cells in cellular automata and tries to modify it in the way that certain selected distribution of scales is achieved.

There are three main methodological approaches that can be chosen to measure the urban space as memory: use of adaptive skeletons, as in (Wilder et al., 2011b), use of the straight skeleton as the approximation of adaptive skeletons or use of propagating wave simulation in cellular automata. Adaptive skeletons may be too difficult to implement from scratch, as authors do not provide code, and they state that it is very slow, several seconds per skeleton, therefore, not suited for massive skeletonization that is needed for this research. The approximation with straight skeleton is a strong candidate for urban shapes, as they are usually straight; moreover, there is a reliable implementation, but it is in very unusual setting, i.e., PostgreSQL database (PostGIS, 2018a). Steep learning curve and unavailability of consultants present a barrier for this approach. Therefore, cellular automata is left, which is even closer to the theoretical side of this research; it was used to simulate the propagating waves in neuroscience, and the shape skeleton is only the approximation of this process. The only downside of using cellular automata could be speed; therefore, this technology should be tested on simpler problem by using the same principles that are needed for the urban shape as memory.

2.1. Propagating wave for urban growth simulation

Propagating waves have been observed in multiple scientific areas, connected to this research: neurology, sociology and psychology, but it was not used in the urban context. There is only one study confirming the skeleton approximation in human cognition of space, and it has already been stated many times that skeleton is a mathematical prediction of propagating wave outcome. In order to increase the basis for this concept use in the urban analysis further, an urban growth simulation model, based on the well-known theoretical principles, will be created. Building simulations with simple rules can help to understand the observed phenomena. In the following experiment, simple rules proposed by Alexander are implemented in the difficulty increasing manner. This approach allows finding the minimal required difficulty of rules to replicate the desired result.

Christopher Alexander's book *A Pattern Language* consists of a list of recurring problems and suggested solutions to them (Alexander, 1977). The book covers different scales, starting from the distribution of cities and gradually reducing the scale down to the room decoration. In the real world, usually, the problems do not come alone; this is as well embedded in the pattern language. The patterns have lists of other linked patterns, which unusually come together; the instructions are presented as well on how to follow those links and choose the right patterns for problems in one’s project. Modern architecture is increasingly made to impress, leaving behind many enduring principles, even the ergonomics; this could by slightly compared to the behaviour known in the animal kingdom and primitive cultures, better known as "who shouts louder". Therefore, the authors hope that the pattern language could become a framework that would help and not fall into the primitive ways by bringing back the traditional and "timeless" solutions to the recurring problems. The authors carefully hedged the sentence about the rules being not constraining: "we have tried to write each solution in a way which imposes nothing on you". Therefore, several patterns are concluded with clear instructions in an inquiring tone, which reminded of the law, and programming was a part of the inspiration for this research as well. It was widely used by architects and designers; surprisingly, it as well inspired the programmers to create an analogical practice for themselves. The software design pattern idea was first presented at OOPLSA in 1996 and originated from Alexander's patterns that describe the common practices in software engineering (Beck, Cunningham, 1987). Interestingly, crypto currency technology "block chain" is actually one of the software design patterns that originated from the same idea (Shvets, 2007).

Alexander's ideas and software design are behavioural patterns with set of rules or requirements describing them, but a more widely used meaning of this word is visual patterns. Although the meaning is different; nevertheless, it is coherent, considering that usually perceived pattern as well follows some rules, recurrently involve repetition. The deriving rules or mathematical descriptions of visual patterns are old and widely known problems; one notable paper is *The Chemical Basis of Morphogenesis* (Turing, 1952), where the analysis of oscillating chemical reactions with resulting visual patterns was predicted. The computer simulation of visual patterns in nature is investigated by the science of pattern formation.

2.1.1. Analysis of the pattern rules

In 1996, in the presentation at the silicon valley, Alexander expressed dissatisfaction with the current situation of his patterns usage and suggested a production of tools, which would help with better assimilation of his patterns

(Archives: Keynote Speech to the 1996 OOPSLA Convention, 1996). It was impossible to find the attempts to simulate, generate or diagnose solutions based on the Alexander's patterns; therefore, this research attempts to fill the gap.

There are three types of patterns: urban, architectural and interior design, as the aim is to extend the toolset available to urbanists, the selection of pattern will be made from the first type. This is the first attempt to simulate patterns; therefore, simplicity and lucidity are among the most important features in the pattern description. Pattern No. 3 has been selected because it has clear rules, and the illustration of the pattern is presented. The simplicity of the pattern is as well a motivation for this choice. Further on, it could serve as the foundation for more elaborate simulation that would include other patterns. There are several urbanistic patterns that intuitively should be compatible and work in synergetic fashion, but it must start somewhere, and pattern No. 3 is the first one from this synergetic group; therefore, it has been chosen. Moreover, in this study, it will play the most important part, considering the fact that it will be used as a visual test to evaluate the result of simulations, not dissimilar to the fitness function.

In pattern No. 3 City Country Fingers, the problem is the overspread fabric of the city, in which people cannot react to nature and feel trapped. This occurs when the distance to the nearest countryside is more than 10 minutes away in a walking distance, i.e., about 1 mile. The suggested solution is to form and maintain bands, or as he calls them "fingers" of urbanized and natural territories. An additional motivation is that the problem as well arises from the conflict of planners versus the private sector. In case the rules of this pattern become the city planning code, it would be exposed to the rule bending attempts from the private sector; therefore, it should be very clear and not cause misinterpretations.

Several other patterns address this problem as well. Pattern No. 2 The Distribution of Towns tries to limit the population of big cities and encourage the growth of smaller ones. Four types of town are introduced based on the population, ranging from 1 million to 1 thousand in decrements of division by 10. The rule is not let the city grow to a bigger type if it is closer than the threshold of the type, which are 250, 80, 25, 8 miles, accordingly. Pattern No. 5 Lace of Country Streets encourages to place streets in a square grid at least 1 mile apart and grow the city along those streets, leaving farm areas inside at least one square mile away. Looking just at the description, one can get an impression that those patterns are not compatible with each other, although later on, there will be seen a slight change of rules that can produce a result that is visually similar to the not intended pattern.

2.1.2. Implementation of methodology

In order to simulate the pattern, python programming language was used (Rossum, Drake, 2011a). The simulation was implemented by using the paradigms of cellular automata; therefore, every plot of the land is treated as a simulation cell. In order to store the parameters of cells array, the arrays of values are defined as global variables for every parameter in the cell. The visualization is done with Pygame module; cells are drawn with "Rect" function, label in the cells with "font" object (Lindstrom et al., 2000). The simulation needs to calculate the shortest path distances; therefore, Dijkstra's algorithm was implemented to calculate them. For this algorithm to work, the graph structure is needed that represents how the cells or nodes in graph structure are connected. Graph links are not stored, rather generated with function on demand. There are 3 types of graph structures implemented, connecting only 4 sides of the cell neighbours, 8 neighbours and 8 neighbours with different costs approximated by Pythagoras theorem for better linear distance simulation. Graph nodes on the edge of the simulation are connected to nodes on the opposite edge; therefore, all nodes have the same number of links, and the edges of simulation do not have special rules, thus prevent skewing of the results.

2.1.3. Adaptation of rules

Some suggestions are provided how to handle this problem on hilly terrains, but the solution is too obvious; therefore, it was assumed that the simulation is on flat terrain.

With no doubt, one could adopt a general plan of the city to look like a pattern in the illustration and use it to process the incoming requests of land use change. It is not known how long the "fingers" should be. It would be preferred to find the simplest set of rules in the simulation that would produce this pattern.

If it is assumed that municipality worker follows the rules of the pattern and issues permits to land use change requests based on pattern No. 3 City Country Fingers, the system is initialized with farmland and convert one cell to a city. The essence of this pattern is the prevention of uncontrollable city growth. It is safe to assume that during the process of city growth, the owners of plots that are located closer to the city centre would want to convert them first, which is as well confirmed in the problem of the pattern.

In every simulation step, the owners of the farm plots next to the city apply for permits to convert the plots to the city. If the city is larger, it is assumed that the most interested would be those who have land closer to the city centre; therefore, they are evaluated first.

The city centre is found by measuring the distance from the city edge, using Bellman–Ford algorithm. For this algorithm to work, the cells are treated as nodes in the graph. There are some choices to be made, which later result in different visual outcomes. For example, the graph is constructed considering only four neighbours (Manhattan distance), diamond shape cites are generated in case nine were connected and square cities were generated (Figure 5).

Manhattan distance is referring to the well-known street grid like topology; therefore, it applies to graphs with horizontal and vertical connections only. Consequently, the distance in such graph is measured not by Pythagoras theorem but by summation of the differences in coordinates from the start to the finish (Black, 1998).

In order to approximate the distance to the next cell and diagonal cell better, the cost of movement was used 2 and 3, respectfully. This is the approximation of true distance 1 and 1.41; then, the circle cities are generated. In the illustration of this pattern, the "finger" endings are rounded; thus, the latest choice of graph generation could be the best choice.

The calculations are performed with processor with built in floating point operations; therefore, much calculation speed will not be gained; nevertheless, the integer numbers are chosen because they fall into simple rules category, are easier to interpret, therefore, enable easier debugging, and the labelling of cells in the value visualization will not require the approximation or conversion of units.

After one plot has been converted, the step is over, and the distances are measured again. This results in the same problem that was described by Alexander, i.e., uncontrollable city growth. The maximum allowed city depth should be set as a means to limit the city growth. The depth is the maximum reachable distance from the city edge. As in the tree graph, the depth is the distance from the trunk to the leaf (Cormen et al., 2009). For 7 cells maximum, the allowed city depth will result in one cell dimension, 115 m is close to 91 m (Lynch, 1960), 100 m is the ideal size for a large square form (Shaftoe, 2012), and 150 m is the ideal urban climate (Ng, 2009). Moreover, the municipality worker of virtual city now has to measure the city depth for every land use change that is requested. This is only possible with running path, finding algorithm on the imagined city with requested converted plot. In order to speed up the process, the "black list" of plots, whose conversion results in deeper city centre than allowed are kept and never evaluated again. This simulation successfully limits the city width; it produces the flowing trail of the deep city plots, corresponding to another Alexander's patterns: ring roads, promenade, shopping street, but results in small farms, the size of one plot only.

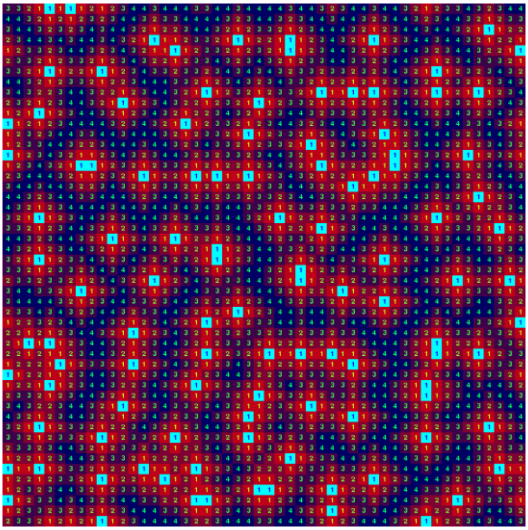


Figure 5. One plot farms, product of simulation, lacking farm size maintenance

Small farms are not effective. If the city border is spiky with narrow bays and peninsulas, it would be ineffective for the farm usage. The concept of grey zones was introduced to increase the farm size. In order to find grey zones path, the finding algorithm was used to find plots of the same depth as the city centre. After that, the farm centres were used to find the shape of the effective farm (**Figure 6**).

City border spikes or peninsulas moved farm centres away from the city centres and caused areas around the spikes or peninsulas and indents or narrow bays not to be covered by effective farms. Those not effective farm plots were marked as the grey zones. They are counted, and this count is used as the evaluation in winner plot selection. The algorithm tested all possible city border expansions in this manner and selected one for conversion, which has the minimum grey area. This was computationally intensive, and the difficulty increased the proportionally with city border length. Despite all that, the effort of this approach did not produce desirable results: more farm plots are generated, but still most of them are one plot size. Increasing city only by one plot in most cases produced grey plots, and even though the minimum one was selected, it was not enough to widen the farms. This approach could be improved by analysing the groups of plot conversion candidates, but then, the computational difficulty would increase exponentially.

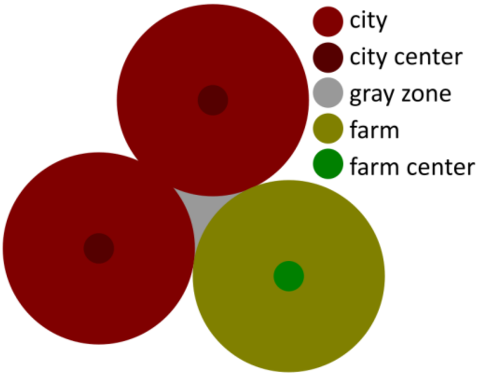


Figure 6. Concept of grey zones, too far from the city centre to be converted to the city, too far from the farm centre to be used as a farm

Instead, the concept of search was changed. The simulation was initialized with wilderness and one used plot. The used plot means that it is not wild, used for a farm or a city. The plot evaluation for changes from “wild” to “used” proceeds in the same manner as in the second model. The assumed distance between the city centre and the farm centre should be the sum of their depth or one mile. The distance to the centre is measured as before, and the assigned functions are as follows: if the plot is closer to the used land’s edge, it is assigned to the farm; if it is deeper than farm’s depth, it is automatically converted to a city; if plot conversion to used results in a city plot with a depth that is bigger than allowed, it is marked as farm centre. This simulation result could be identified as the pattern Lace of Country Streets, except the city part is wider as described in City Country Fingers. In Lace of Country Streets, it is suggested that the territory that is assigned to farms should be rectangular; however, the simulation generates circular farms; this is caused by the selection of graph generation type, as discussed above (Figure 7).

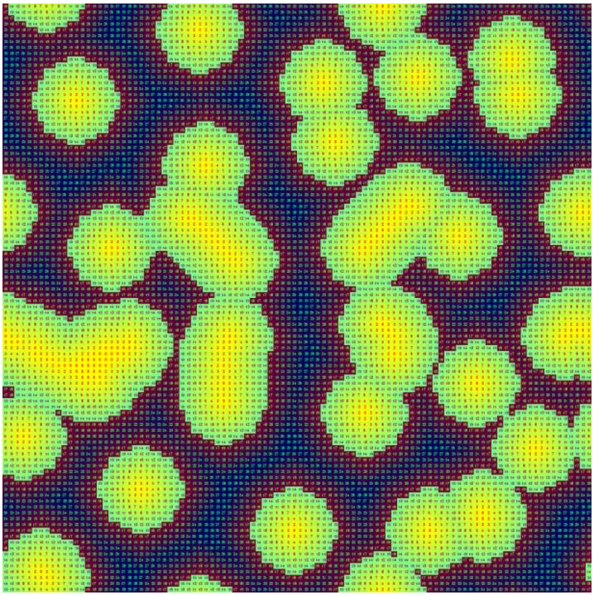


Figure 7. Maintaining farm size is not enough

In the description of City Country Fingers pattern, Alexander writes: "keep interlocking fingers", but in last simulations, the city cuts of farm fingers. It has been decided to increase the difficulty a little in order to preserve the continuity of farm fingers. The algorithm checks all neighbours of the farm centres, if only one is left wild, then it converts it to the farm centre (Figure 8).

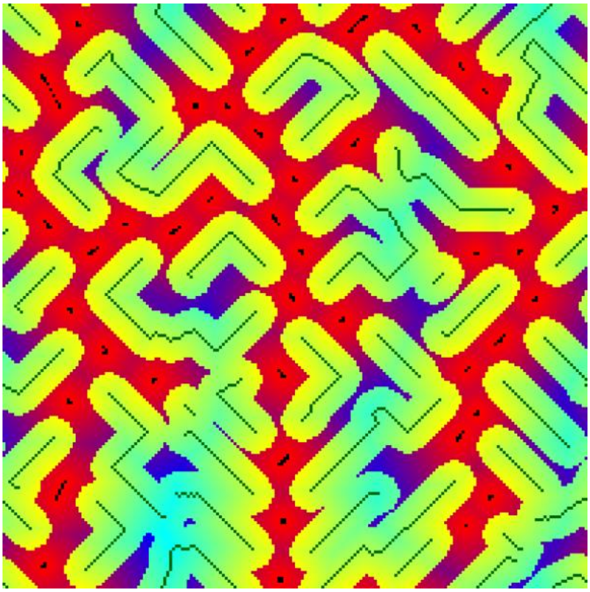


Figure 8. Continuous farm centres

This could be improved further by using fuzzy location of the city centre, as in fuzzy logic, where intermediate values between true or false can exist (Novák et al., 1999). Now, the centre is considered only the plot, which has certain depth; whereas in reality, there are seen rough cliff transitions from suburb to centres. The transition occurs gradually; there would be some plots, which would serve partially as both city and the city centre. Farm centre generations are almost always locked in diagonal direction. This is caused by the rigidity in the grids; this could be improved as well by introducing fuzzy logic, although an increase of radius of the farm continuity should improve the farm's continuity direction. Although the simulation was getting closer to generating visually similar pattern from the book, the rules of simulations were becoming too difficult. Fuzzy logic and city centrality amount of the plot are not simple concepts.

2.1.4. Tuning reaction diffusion simulation to achieve the visual similarity

The reaction diffusion simulation generates visually similar patterns (Pearson, 1993). However, this behaviour is emergent in a very narrow area of 4dimensional space of model hyper parameters, changing the hyper parameter values does not allow tuning the pattern size in cells, therefore cannot be used as a city simulation tool. It is safe to conclude that this pattern cannot be simulated by using simple rules. During the research of previously accomplished pattern simulations, a pattern with visual resemblance to the illustration of the City Country Fingers pattern was found. The reaction-diffusion model in a very narrow parameter space has a visual resemblance to City Country Fingers pattern; therefore, it is considered as a possible alternative urban simulation, if tuning of parameters will allow customising it.

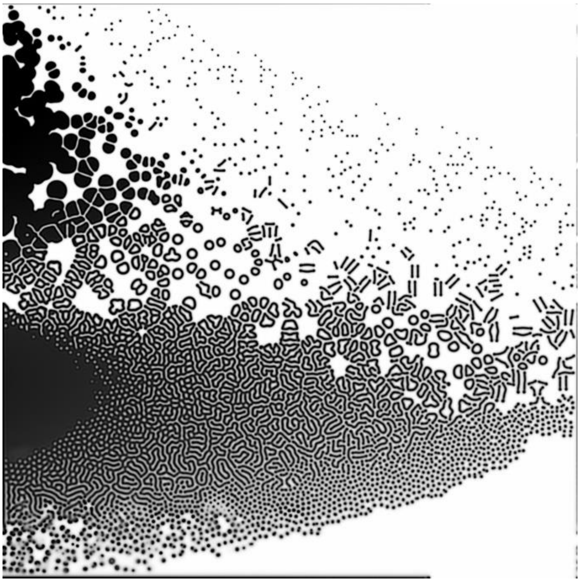


Figure 9. Reaction-diffusion model parameter scan: Diffusion A=1.0; Diffusion B=0.5; feed of A=scan on X-axis from 0.01 to 0.115; kill of B=scan on Y-axis from killMin to 0.07; killMin= scan on X-axis from 0.05 to 0.062 (in the middle) and back to 0.05

This is the parameter scan of reaction-diffusion model (Figure 9). Certain configuration of parameters produce pattern that is visually similar to City Country Fingers, urban planning pattern from Christopher Alexander's book *A Pattern Language*. It offers various other patterns; they as well differ greatly in the temporal domain. However, there was found one similar to City Country Fingers and static convergence. However, it is not usable, because only one scale is measured in cells. All available scans were performed with the conclusion that in other available parameters, there are only very narrow domains where patterns appear and scanning that does not allow scaling the pattern to wanted dimensions.

Figure 10. Reaction-diffusion model parameter scan: Diffusion A=scan on X-axis 0.6 to 1.1; Diffusion B=scan on Y-axis 0.25 to 0.75; feed of A=0.029; kill of B=0.057

Although this simulation is able to produce a more visually similar pattern, it is only available in a very narrow domain of multidimensional parameter space of a model. Even a slight adjustment of parameters destroys the pattern, and this makes it a hypersensitive and unusable model for the research needs.

With the first model, it was possible to achieve results with partial similarity; it generates features that are described in the book; however, the features are not exactly the same shape. For now, it seems that the binary states of the cell are the reason for the lack of visual similarity. This theory could be tested in the future. For now, it cannot be called a complete success, but it shows a potential of such approach.

In the process of trying to implement simple rules, the propagation pattern from axial curves was accidentally implemented. The propagation from city centres reserve land for the city and later, in bigger distances, for the land. Farm centres are formed in the cell where two propagating waves meet. It seems that this implementation of Alexander's pattern is similar to the skeleton formation with cellular automata.

2.2. Search for universal distribution

Golden section, Fibonacci sequence, universal scale and distribution as well as fractal number, they all related to the proportions and were proved to have an impact on the aesthetic feeling, calmness and health, therefore, on cognition of environment, and could be an important factor for urban space as memory. Although there are plenty of confirmations that this is true, some ideas are not tested. The proportions can be measured with methods of computational aesthetics, similarly like it is done in the fractal analysis by calculating the angle of linear function in log-log plot of measurements; therefore, making tiny adjustments to the picture that only increase the selected measurement of the same picture model would produce something recognisable, hopefully a fractal.

Humans are capable to distinguish different types of noises. The example types are white, pink, brown. If analysed on the frequency scale, it turns out that different types of noise are in fact different distributions of frequencies. For example, the white noise would be a uniform distribution; pink and brown are both power-laws of different order (Schroeder, 1991). It is as well quite easy to distinguish distribution of random circle sizes. Consider two pictures, they both have the same number of circles and the same area occupied by the circles, but the distribution of sizes in the first is uniform, and in the second, it follows the power-law (Figure 11).

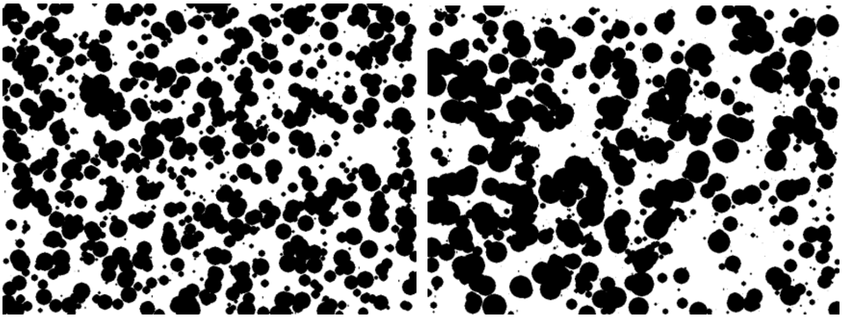


Figure 11. Random circles with uniform (on the left) and power-law (on the right) distributions of sizes

In the second picture, the large circles are rare; therefore, they have to become bigger to be able to occupy the same area.

If considering an image, which starts as white, then random pixels are turned on and off. Sometimes, those random pixels would be adjacent and form accidental lines. It is possible to scan through such image and count the length of lines. Even though the random function that is used has uniform distribution, the resulting distribution of sizes is close to power-law distribution. One of the possible statistical measurements would be to test if the resulting pattern detail sizes are distributed under the power law. It is possible to measure the distribution of lengths; therefore, it is possible to calculate the difference from the desired distribution or in other terms, error. Using such error function, it is possible to perform simulated annealing, or in simpler terms, random search to get the configuration closer to the desired distribution. It is achieved by flipping the value of random pixel and calculating the distribution of line lengths and error function, and only accepting the changes that do not increase the error. As starting configuration was close to the power-law distribution, the accepting of the changes that minimise the error of the power-law distribution quickly converges, but the changes are so few that it they not noticeable in the image.

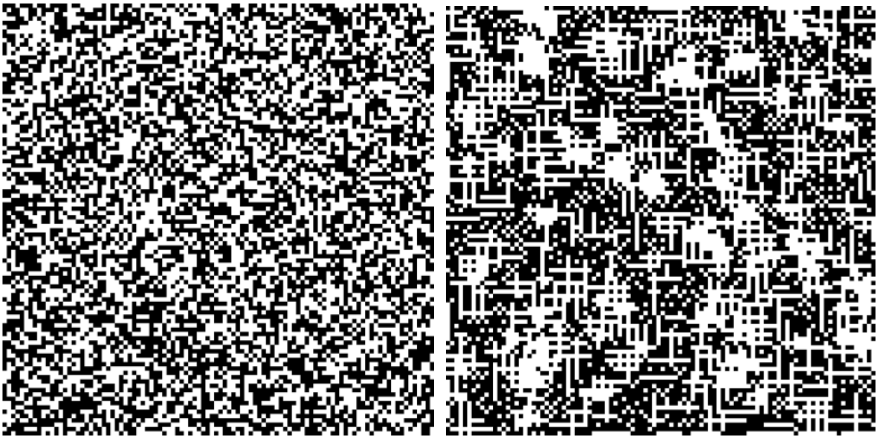


Figure 12. Accidental lines in random picture with length distribution of power-law (on the left), uniform (on the right)

If the uniform distribution is used, it takes quite a while for it to converge, and the resulting image exhibits noticeable structure. Noticeable self-similarities suggest that it could be categorised as fractal or misidentified or a result of cellular automata, although cellular automata rules were not used to generate it. Cellular automata concentrate on the local rules to produce an emergent result; this experiment is exactly the opposite: it concentrates on the evaluation function. This is a stress test power-law as the evaluation function (**Figure 12**).

The unexpected result shows that the power law is not enough; depending on the problem and measurement type, other distributions could be important as well. In fact, the importance of the power-law was criticized, stating that it is an artefact. Brownian motions can produce power laws in different measurements (Touboul, Destexhe, 2010). Newman has collected a wide array of naturally occurring power laws (Newman, 2013). Power law describes distribution; sometimes, it is even referred as a power law distribution. It is close to Logarithmic distribution, which was discovered before the power law to be useful in describing certain phenomena occurring in nature. Before the logarithmic distribution, there was the Gaussian distribution, often refereed as the bell curve, which was used to describe many observations in different fields. It seems that there is certain fashion wave after someone uses certain distribution for the deception of observed phenomena; others use the same distribution as well. After the exhaustion of available phenomena that could be described by the same distribution, another distribution comes into fashion, and it continues. For now, the uniform distribution is only noticed in distances between the nearest members in biology; on the contrary, it is not noticed in scale; therefore, it could be the next fashionable distribution. Despite the fact that this test model produced the opposite result to the one that is predicted by the theory, it still showed that testing for distribution is significant; therefore, the distribution test should be used on the skeleton graph analysis to measure the parameters of urban space as memory.

2.3. Investigation of urban space as memory influence on the social capital

With term modelling, it is referred to Machine Learning models popularly, better known as Artificial Intelligence. “Machine learning: Field of study that gives computers the ability to learn without being explicitly programmed” (Samuel, 1959) is a citation from a famous paper about the early artificial intelligence model that could learn the rules of a board game, i.e., checkers, by playing it with itself so well that it defeated its creator. With this example, two types of machine learning should be discussed; there are more, but only two are needed to communicate the idea in mind. The reinforcement learning, as in the example with checkers, is where the winning strategy is unknown and there were no data recordings of the humans playing. Another is supervised learning used when data is available; it is usually used to solve problems that are easily solvable by humans, for example, recognizing handwritten digits, but are difficult to program explicitly. The problem of this research, i.e., prediction of social capital from the urban form as memory data, is philosophically somewhere in between the supervised and reinforced learning, because it is only probable from the literature research, from which this model can be made, but it cannot be verified with human expertise because such expertise does not exist. It is possible to acquire data; therefore, technically, it would be supervised learning. In order to model urban form or anything (in that matter), using supervised learning data, a matrix is needed. It has already been shown that straight skeleton is most mathematically pure methodology of shape parameterisation that has evolved from Blum's median axis, synonymous with Leyton's process grammar for shape and shape as memory. It has already been discussed how Wilder analysed the graph structure of skeletons to obtain a data matrix (Wilder et al., 2011c).

Most of machine learning models require data to be prepared in the data matrix. The models, which do not have such requirements, are usually specialised for a particular type of data, which is incompatible with data matrix, some examples include: images, text, sound recordings.

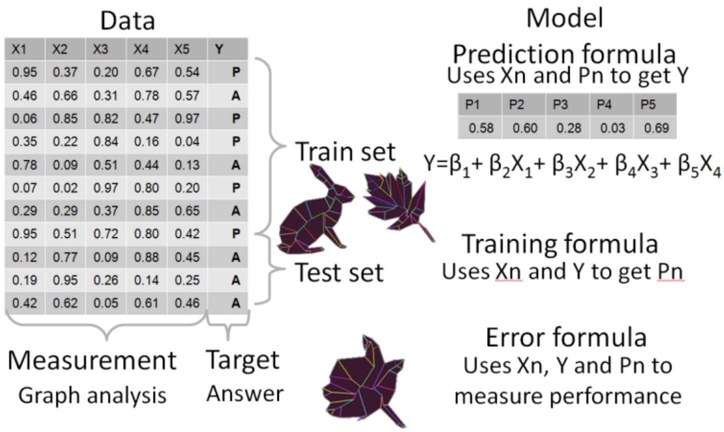


Figure 13. Basic example of a machine learning model (multidimensional linear regression)

Most machine learning models try to find the parameters of selected mathematical function, constants of a formula, in such way that the formula would represent the data as good as possible. The goodness is often measured by the mean square error formula:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y_i})^2;$$

where MSE is the mean square error, n – sample size, i=1 – summation starts from the beginning of the sample, Y_i – known answer, $\widehat{Y_i}$ – prediction of the model. The difference between the actual data and the prediction $Y_i - \widehat{Y_i}$ is squared to be independent from the ring and emphasise large errors against small. This is repeated for each data point, and the average of all squared errors is calculated. MSE function represents a multidimensional field of errors where the solution is the lowest error value. The models, using this formula, can be named the least squares regression. In this case, the regression represents a strategy of error reduction, which is done by another formula used to calculate how to change the parameters of the function. This is done by calculating the slope of error at the current point. When the lowest point of the error is reached, the model is trained, and the parameters of the function represent abstracted knowledge that is gained from the data. The parameters of some models, like linear regression in this example, can be understood and interpreted by humans, but in more elaborate models, this is usually impossible; in those cases, other statistics are calculated, which help to understand the inner workings of the model. However, in most cases, the usefulness of the model is associated with the accuracy of prediction. Using standardised statistics to measure the model accuracy is a reliable factor of comparison and evaluation. One of the most common statistic accuracy is the coefficient of determination (full description is provided in terms and definitions). Measuring the accuracy of the same data, which was used to train the model, would not be honest and is considered an error. Therefore, Cross Validation strategy is recommended to overcome this. Using this strategy, the data is split into at least two parts: the part where the model is trained, named train set, and the part that is not used in training, but used to measure the accuracy, named test set. The same strategy is planned in this research. Further on, it is planned to test and compare many models.

Putnam's methodology for social capital index calculation already produces a data matrix, although the scale is different, it produces one index per administration unit, for example, county, but when using the state of the art skeleton technology, it is impossible to get a skeleton of such a big territory. This situation should be analysed in order to find a solution for this limitation.

In order to analyse the urban shape as memory, it is possible to generate a straight skeleton for every building, but for the spaces between the buildings, some other strategy is needed. It was chosen to split the area in a regular square grid, which could be used to cut the building out of it. The resulting grid with the cut of buildings could represent a space between the buildings; moreover, it would be possible to generate straight skeletons for them. There was chosen 300 m as the dimension of the square grid by analysing the literature sources. Alexander states that people cannot identify a neighbourhood that is bigger than 300 m (Alexander et al., 1977). Moughtin noticed that the pedestrian route could be attractive only if it has some interesting object in at least every 300 m (Moughtin, MRTPI, MA, PhD, & BSc, 1999). Jan Gehl research showed that humans can distinguish humans from other animals up to the range of 300–500 m (Gehl, Rogers, 2010). Knox measured that the average pedestrian passage is between 150–450 m (Knox, 2010). The distance of 300 m is compatible with all of these theories and research results, therefore can be used to generate the reference grid.

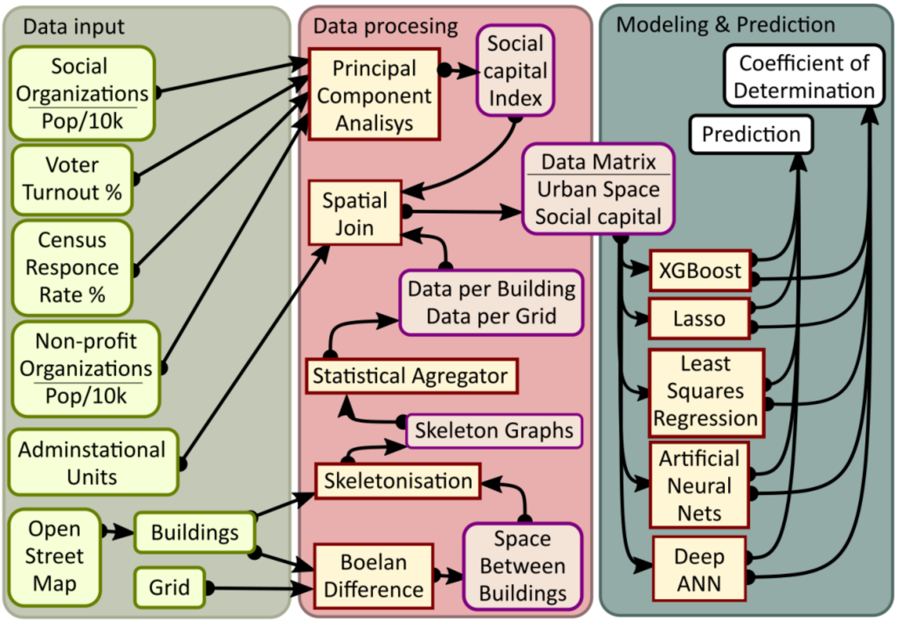


Figure 14. Methodological overview map

In order to generate straight skeletons of a building and the space between the buildings, the data from the prominent example of volunteered geographic information source OpenStreetMap was used (OpenStreetMap, 2019). One of the most reliable and fast skeleton generators is the straight skeleton implementation for the Geographical Information System adaptation plug-in for the PostGreSQL database, accessed with "ST_StraightSkeleton" query (PostGIS, 2018b).

2.3.1. Straight skeleton preparation methodology for urban shape as memory analysis

In order to be able to use the social capital database together with the results of building shape analysis, the buildings had to be analysed in the same territories or counties. The US Census Bureau provides Congressional District Cartographic Boundary Shapefiles (US Census Bureau, 2016). They provide GIS data projected in the North America NAD83 projection and World Geodetic System WGS84 projection (OpenStreetMap, 2019). In order to achieve compatibility, the county boundary data was re-projected by using the open source software package (QGIS, 2017a). OpenStreetMap data was not downloaded directly from the service; instead, the North American data extract in one file was used, kindly provided by (Geofabrik, 2016).

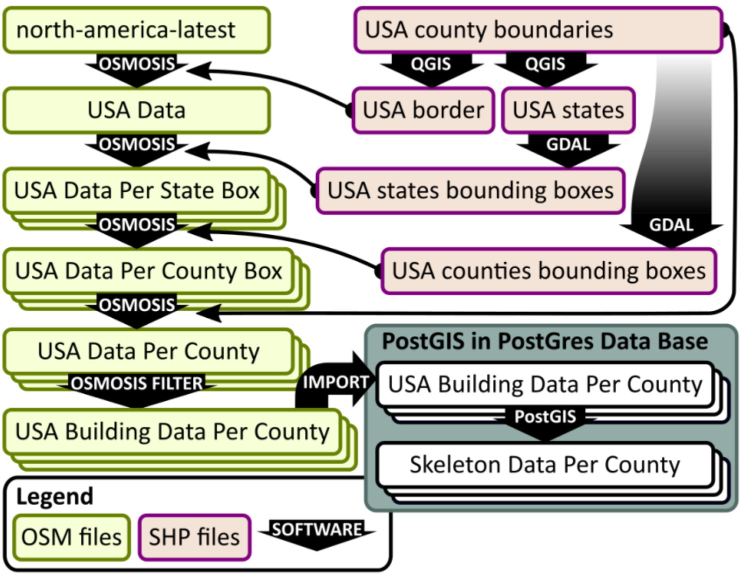


Figure 15. Data preparing pipeline

Cutting into county boundaries and filtering building data was done with an open source software tool (Osmosis, 2016). Due to the large size of the data file (169GB), the cutting was very slow. In order to increase the speed, the pre-cuts were made into a bounding box, covering all the state boundaries. In order to get the state boundaries, a dissolve tool was used in the QGIS. A special script was written by using the OSGeo library in Python environment to get the bounding box coordinates

(OSGeo, 2016; Rossum, Drake, 2011a). Osmosis needs a special “.poly” file format to be able to cut by polygon. The county boundaries were exported to poly, using a specialised plug-in in QGIS. A skeleton generation algorithm that was used is custom and not available (Wilder et al., 2011c). Judging from the description given by Feldman (Feldman, Singh, 2006), it would be computationally too expensive in this application (Feldman, Singh, 2006). Therefore, for the skeleton computation SFCGAL, an extension to PostGIS was chosen (SFCGAL, 2016; PostGIS, 2018b). The skeleton implementation in SFCGAL is based on (Aichholzer et al., 1996). Osmosis has integration into PostGIS, which was used to upload the building polygons into the PostGIS database. In order to compute the skeletons, the "ST_StraightSkeleton" function of SFCGAL in PostGIS was used. Following the structure of the data from the files to the database tables, at this point, one table in the database holds information of all the building skeletons of one county (Figure 16).

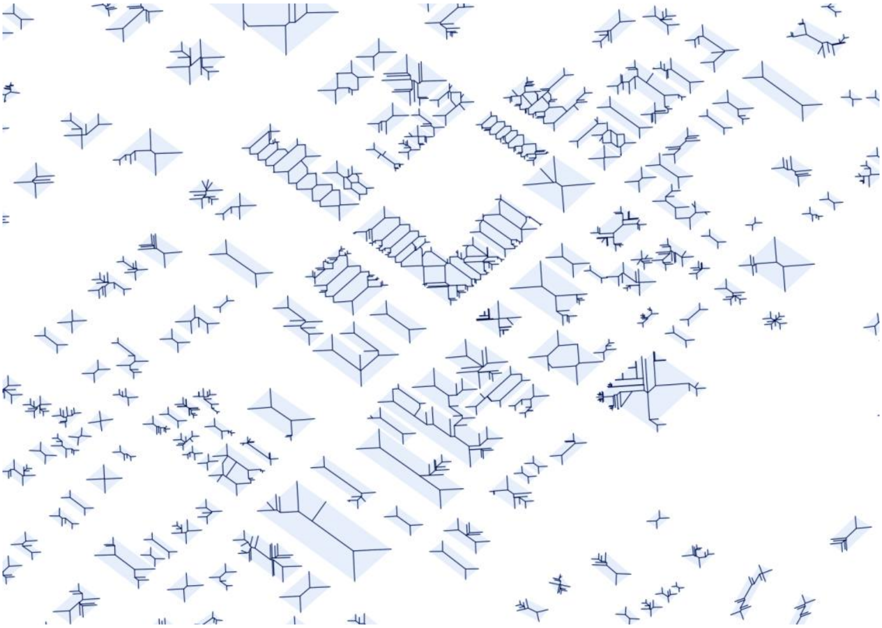


Figure 16. Building polygons and skeletons

Therefore, there are as many tables as there are counties, and every entry of the table consists of graphical representations of a skeleton that should be analysed to extract urban shape as memory parameters.

2.3.2. Straight skeleton analysis methodology for urban shape as memory parameter extraction

The information about the skeleton nodes is kept inside a single cell in the database entry. A special function in PostGIS was developed to explode a single skeleton into a table, consisting of entry per node, and extract the analytical parameters. The parameters described by Wilder can be summarised as a network analysis of a skeleton and does not have open source libraries to compute (Wilder et al., 2011c).

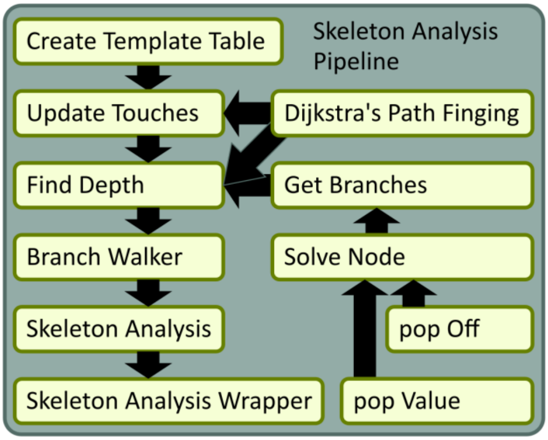


Figure 17. Skeleton analysis pipeline

Several customised variants of the Dijkstra's algorithm, path finding algorithm, were implemented, which collect all the parameters along its execution (Dijkstra, 1959). It firstly initiates a depth parameter with a number of nodes in the skeleton, and then starts at multiple nodes, which are dead ends of the skeleton network and follows to find the deepest node in the skeleton, writing steps taken from the dead end into a depth parameter, in case it has a bigger value than the steps that were taken. From this, it is possible to calculate the maximum and mean depth of the skeleton. Then, it starts at the deepest node and steps through the nodes by calculating the number of branches it encounters and the angles between nodes. From these values, it is possible to calculate the maximums and means of the following parameters: the number of nodes connected to a given node, the length of the node, the azimuth of the node, the turn angle between the nodes, the absolute turn angle between the nodes and the branch depth as well as the total number of branches and “wiggliness” of the skeleton. Later, the dividing sum of all turn angles is calculated from the sum of the absolute turn angles.

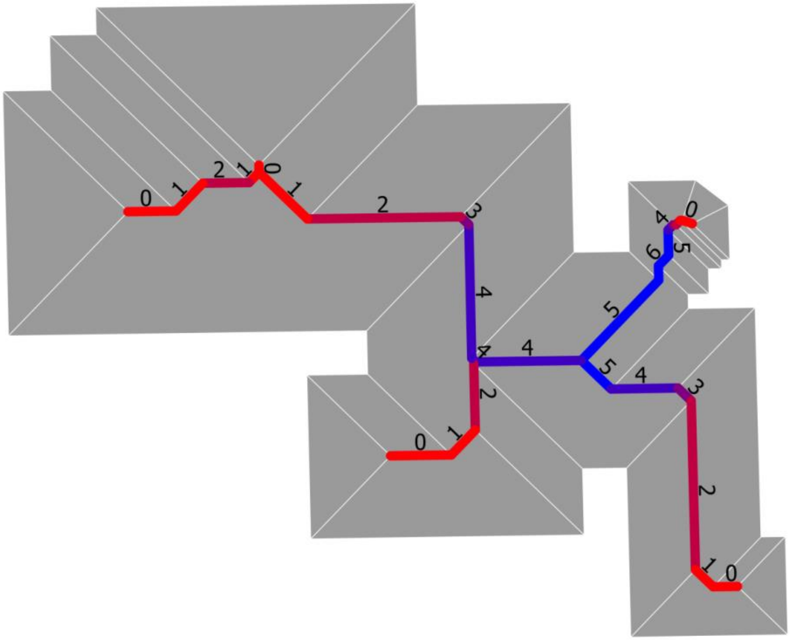


Figure 18. Pruned skeleton, (red and blue) colours and numbers represent the skeleton depth

The skeleton generation algorithm used by Wilder allows skeleton optimization, whereas the SFCGAL implementation that is used in this study does not; furthermore, it does not offer any tuneable parameters. In order to compensate, the approach proposed by Bai was used (Bai et al., 2007). The skeleton nodes with minimum depth were dropped, and the calculations were repeated by collecting the parameters in a separate set. The pruned skeletons have a deeper topology and no spurious branches (Figure 18). A full algorithm is provided in the supplementary materials.

At this point, there is a collection of parameters per building, although in order to connect to social capital data, it is necessary to have parameters per county. In order to reduce the entries, a bucket sort algorithm is used (Cormen et al., 2009). It increases the number of parameters based on the maximum value. The resulting database consisted of 116 parameters.

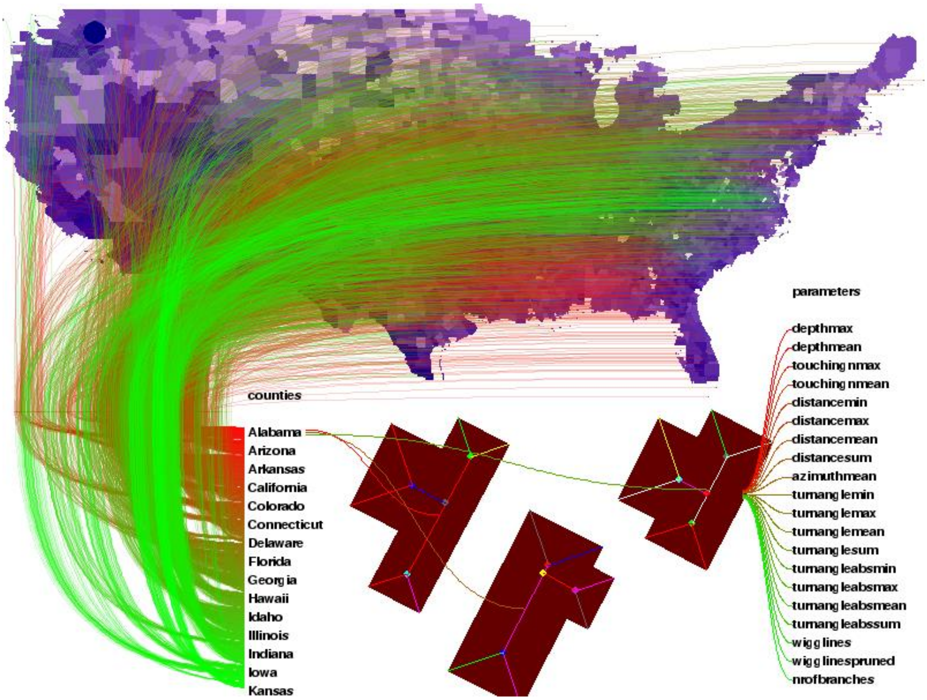


Figure 19. Parameter aggregation from the skeleton graph analysis (lower right) to the aggregated and geoposition parameters of counties (the upper part)

For the analysis of the data, the open source package Scikit-learn was used (Cournapeau, 2018a). Naturally, there were many small buildings with simple shapes and progressively less, as the shape complexity increased. Salingaros and West argue that the environment, which does not follow the rules of universal scale and universal distribution, negatively affects the health (Salingaros, West, 1999). These rules are better known as the power law (Clauset et al., 2009). The question is whether it could affect the social capital as well. In order to test this theory, a collection of parameters that were generated with a bucket sort algorithm were fitted with a toolbox for testing if the probability distribution fits the power law (Alstott et al. 2013). This fitting was done in every county per parameter. The fitted models produce four additional parameters, representing how well the parameter distribution fits the power law. The resulting database consisted of 2400 samples that were obtained with 880 parameters.

In order to show a connection between the data parameters, one can tune weights and values of the machine learning model by exposing it to the data. The ability to represent the data of the fine-tuned model can be measured by calculating the difference between the data and prediction as well referred as an error. Several models fitted for the test if it is possible to predict the social capital from a statistical analysis of building skeletons. In order to measure the model performance, the R-squared method was used (Colin et al., 1997). The most simple and best-known model, the least squares regression was used as a benchmark for other models

(Geladi, Kowalski, 1986). It was impossible to use the naive Bayes classifier as in (Wilder et al., 2011c) due to the incompatibility of the problem type, they had a classification problem, in this research, the continuous value is attempted to predict. Thus, the Bayesian ridge model was used, as it is based on the same theory (Park, Casella, 2008). Lastly, a multi-layer perceptron regressor was used (Rumelhart et al., 1985). A combination of hyperbolic activation function and quasi-Newton solver showed the best results. It was possible to increase the predictions further by eliminating some parameters with a forward selection algorithm (Koller, Sahami, 1996).

2.4. Social capital of the United States of America

For social capital data, a readily available database was used (Rupasingha et al., 2006).

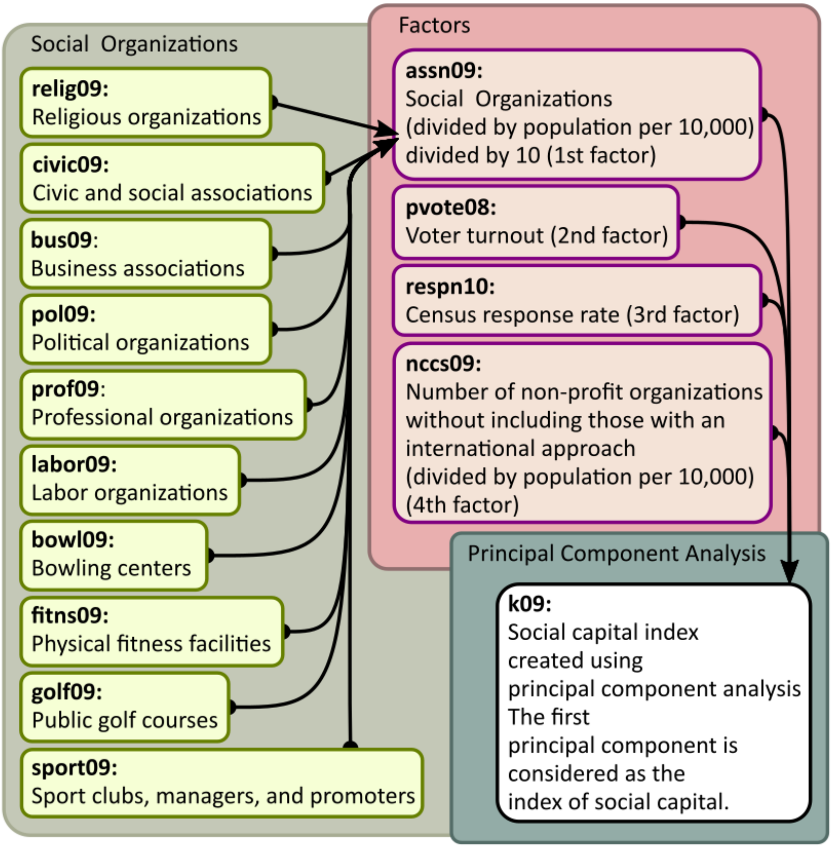


Figure 20. Social capital calculation methodology based on (Rupasingha et al., 2006)

There, the social capital follows Putnam's theory (Putnam, 2001a). The database consists of input variables, including the numbers of religious organisations, civic and social associations, business associations, political organisations, professional organisations, labour organisations, bowling centres, physical fitness facilities, public golf courses, sports clubs, population, voter turnout, census response rate, number of non-profit organisations, not including those with an international approach, and the final computed value of the social capital index. It covers the states of the USA with the exception of Alaska. It has an entry for every county in each available state, in total 3,108 entries (Figure 21). This database defined the territorial area and the structure of the research.

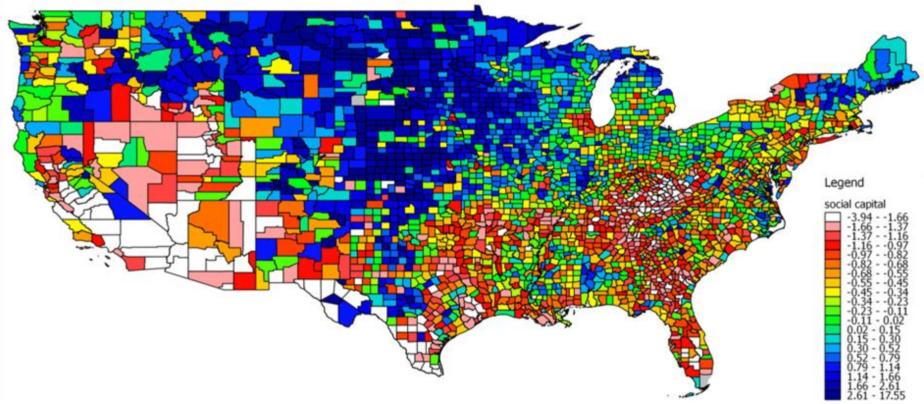


Figure 21. Social capital of the USA counties, data from (Rupasingha et al., 2006)

2.5. Urban shape (as memory) connection to the social capital of Lithuania

In order to adapt the methodology to a different location, it had to be adjusted to the local peculiarities.

The graph analysis was made for each skeleton of the building and the space between the buildings. For each arc of each graph, the following were calculated: branching, branch depth (how many branches are already branched), arc depth, arc turn angle, absolute value of arc turn angle, arc angle to north, arc length.

The buildings can vary in plan configuration and the space between the buildings even more; therefore, a graph can have different number of edges; to find out the values of each graph, the values of the edges were aggregated by computing their statistics: minimum, maximum, median, mean, standardized deviation and variation.

Moreover, a single administration unit may have a different number of buildings and their graphs as well as cells representing the space between buildings and their graphs. The administration unit is the smaller unit or subdivision of the municipality; in the USA, it is named county, direct translation from Lithuanian would be “eldership”, English versions of municipality pages use “neighbourhood”, but those terms are confusing, therefore will not be used. In order to get variables for the administrative units, they were aggregated once again with the same functions as the aggregation of graph arcs.

The database communication was handled with Python plug-in Psycopg (Psycopg, 2018).

The statistical aggregation was performed by using Python scientific computation plug-in NumPy (NumPy, 2018). The data joins, based on the spatial relationship, were performed with Python GIS plug-in GeoPandas (GeoPandas, 2018).

2.6. Social capital of Lithuania

In order to measure the connection of Lithuanian social capital with urban form, it is necessary to compile a data matrix, where the data entries would represent the same territories, i.e., Lithuanian municipalities. In this data matrix, the skeleton analysis results are treated as the known vector of multiple values, and the social capital index is treated as the unknown scalar value. Therefore, the compiled models, some of which could be expressed as equations, would predict the social capital index by evaluating the urban form as the variables of the straight skeleton analysis. There were no previous researches performed in Lithuania with the chosen methods. As it was already discussed, Putnam's method to obtain social capital index value is used to execute the principal component analysis of the social factors, such as civic engagement, voter turnout, census response rate and non-profit organisations.

Civic engagement is calculated by adding religious organisations, civic and social associations, business associations, political organizations, professional organizations, labour organizations, bowling centres, fitness and recreational sports centres, golf courses and country clubs, sports teams and clubs, and dividing them by 10, for the population of 10000. It is the most difficult factor to replicate because in Lithuania, such classification does not exist. Moreover, it is difficult to comprehend why bowling is separated in its own category. Intuitively, it does not make any difference; the same conclusion was made after a couple of toy calculations were made. After all, when it is divided by the number of categories, it becomes the average number of institutions per category; therefore, it does not make any mathematical difference in which category it was. Presumably, it was done to relate to the title of the book *Bowling Alone*. This is an additional reason to substitute this parameter.

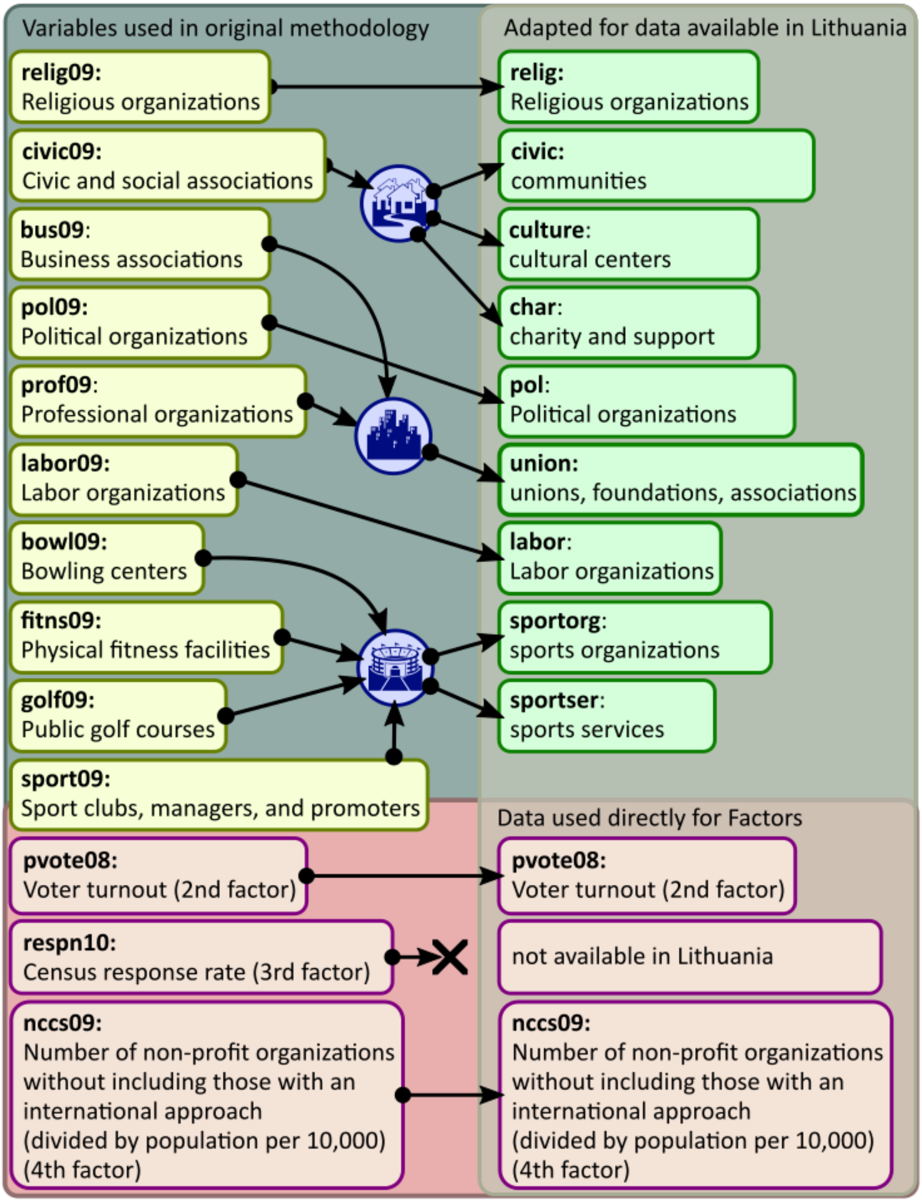


Figure 22. Adaptation of social capital index methodology in Lithuania

The most detailed classification database that is available in Lithuania has 286 classes ("Rekvizitai.lt," 2019). The closest ones to Putnam's method were religious organisations, communities, cultural centres, charity support, unions, foundations, associations, social services, sports organizations, sport services, political organizations. As there were only 8 categories, the formula was corrected accordingly. In order to calculate the non-profit organizations, the factor database provided by info.lt was used (HNIT-Baltic Duomenys, 2015). The data for census response rate factor was impossible to obtain, despite that, according to the methodology of the Lithuanian Department of Statistics, it is necessary to reach 94% of the population; unfortunately, the response to the questionnaire was lower, and the rest of the data was ordered from the State Enterprise Centre of Registers. According to the same methodology, the data on what proportion is answered is not saved (stat.gov.lt, 2011). Therefore, this factor cannot be used to determine the Lithuanian social capital, and there is no data available for it. Despite that, it should not be a problem because the principal component analysis can be made on any number of factors. In order to calculate the voting activity factor, the data provided by the Central Electoral Commission of the Republic of Lithuania for the 2016 parliamentary elections was used (vrk.lt, 2016). The data was aggregated, using the "Join attributes by location" feature of Qgis geographic information package (QGIS, 2017b). The principal component was calculated by using machine learning package "sklearn" in the Python programming language (Rossum, Drake, 2011b; Cournapeau, 2018b).

2.7. Results of modelling urban shape (as memory) influence of the social capital in Lithuania

The variables of building and the space between buildings after the double aggregation that constituted together consisted of a data matrix of 1750 variables and 60 samples, because there are 60 municipalities in Lithuania.

Various models were used for modelling social capital with Lithuanian data only: linear regression, lasso, random forest gradient boosting models LGBM and XGBoost (Chen, 2016). In order to verify the accuracy of the model, a random 20% of the data was excluded to a validation set, which was not shown in the model during the training. During the validation stage, the accuracy was calculated from the validation set.

Table 1. Accuracy of different models

Model	Coefficient of determination (R^2)
Linear Regression	-24
lasso	-6
lgbm	0.1
xgboost	0.3

The best model XGBoost predicts with 30% accuracy. The LGBM model predicts with only 10% accuracy. Linear regression, as well as lasso models, was validated with negative value of R2. R2 is the coefficient of determination. Its calculation compares the errors of a model to the imagined benchmark model, which is the average of all answers and does not depend on the actual input. If a model would predict the data perfectly, R2 would be equal to 1, but if the average error of the model prediction was equal to the benchmark model's error, it would be equal to 0; moreover, if the average error of the model prediction was larger than the benchmark model's error, it would be negative. Positive R2 values can be treated as fractions and converted to the percentage.

Based on the coefficient of determination, there is a statistically weak relationship between the urban form and social capital.

When interpreting the map of Lithuanian social capital (Figure 23), there can be recognized several known trends, like urban sprawl in the ring type administrative units. The data for this map is provided in the supplementary materials.

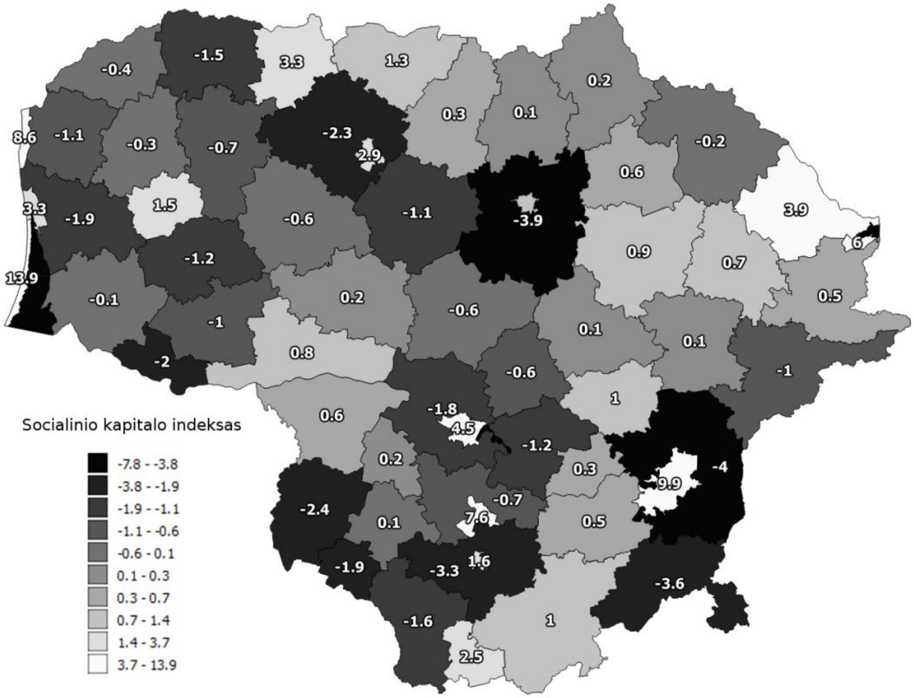


Figure 24. Map of Lithuanian social capital index

Circular municipalities surround the core cities; they offer lower taxes, ability to convert farm land to low density residential urban plots, amplified by human nature to own a fortress, all synergy emerges in the urban sprawl of non-functional low-density residential territories. They become food deserts, i.e., the areas where it is impossible to buy food by walking the distance or driving 10 minutes; nevertheless, all other functions, like recreation and work, are neglected as well. In order to fulfil all their needs, the residents of such neighbourhoods travel long distances and induce load on the transport system, starting a cascade of multiple unnecessary negative factors, like pollution, carbon dioxide emission and energy consumption. Further on, the trend in daily transportation creates potential to develop commercial purposes in the cultural infrastructure in the core municipality.

It is necessary to point out that not all ring municipalities undergo only on urban sprawl; obviously not, there are hotspots of social capital ant culture as well. Nevertheless, the trend is well known, noticeable, and now, as a result of this research, is measurable as well. The largest difference in the Lithuanian social capital index is in Vilnius City and Vilnius District Municipalities. It is possible to notice it in the urban genotype and phenotype traces.

The ring municipalities overall lack infrastructure, but the opposite is true in the resorts. However, it would be a mistake to say that the social capital index of such territories is incorrect; it is merely episodic. It would be safe to assume that this effect is explicit in Nida Municipality.

There are some cities where infrastructure is created by the influence of ring municipality and excess infrastructure for being a resort. It is noticeable in Birštonas Municipality; a synergy between the two trends that were discussed is both a resort and a central municipality with its own ring.

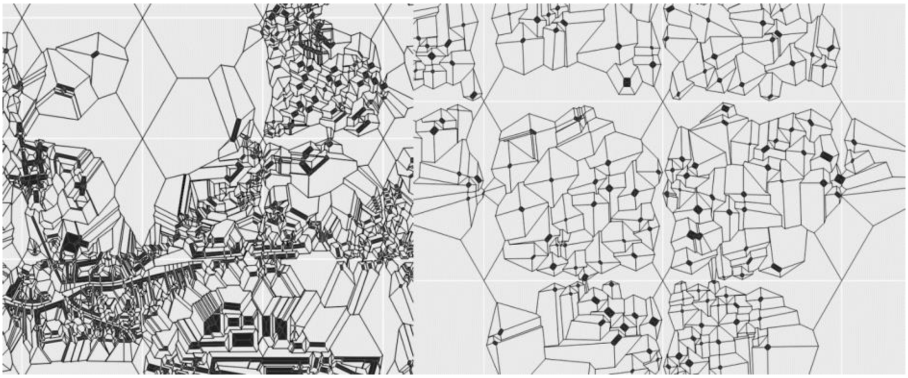


Figure 25. Skeletons of buildings in Vilnius (on the left, social capital index 9.9) and Vilnius district (on the right, social capital index -4)

It is possible to gain intuition while observing the associations of certain urban phenotypes and genotypes with low or high social capital. Putnam's methodology, combined with the specifics of the data that is available, tends to measure the social capital within administrative units, but within the administrative unit itself, the social capital may be heterogeneous. Perhaps, in the future, more detailed and abundant data will be available that will distinguish the boundaries of areas with homogeneous social capital and the focal points of the social capital.

Table 2. Accuracy of different models

Model	Coefficient of determination (R^2)
Linear Regression	-24
lasso	-6
LGBM	0.1
XGBoost	0.3

Random forest gradient boosting models are used with the hypothesis that it is impossible to model a data matrix with mathematical functions. In order to prove that such model was needed, it was necessary to disregard linear models. For that purpose, linear regression and lasso models were tested and converged with a very low coefficient of determination, at least partially confirming this hypothesis. The random forest model attempts to convert each "x" to a discrete variable by finding a boundary (in the value of continuous variable) that it uses to classify the continuous variable. Due to the fact that the modelling of such a new discrete variable results in improved prediction, it is accepted in the model, and the next threshold value is searched, which will add additional categories to the discrete variable and is likely to improve the prediction capabilities. This is the depth of a random tree. The same process is repeated for each variable. Because such a model attempts to categorize and test the improvement in prediction for each variable individually, it does not have a dimensionality problem; in addition, it allows evaluating the impact of each variable to model the performance. Random forests with trees of 4 steps depth were used in this study. Gradient boosting models try to model the error of the previous model in order to be able to compensate or negate it. This allows many poorly predictive models to form one stronger predictor. In this study, the best result was achieved by using 10,000 steps of gradient boosting.

The linear regression model tests the hypothesis that the data matrix has a linear relationship between all x variables, in this study, the results of the urban form the linear skeleton analysis and the y, the social capital index. The parameters of the modular linear function are picked automatically to determine their height at the start of the coordinate system and the angle dependence on the "x" variables.

Such a model suffers greatly from the dimensionality problem. It can be compared to the solution of the system of the linear equations; there are more equations than the unknowns. The answer is found quickly, but because data is not a linear solution, it is inferior.

This is confirmed by the difference in the coefficients of determination between the training and testing samples. The coefficient of determination is 0.79 during training, but drops to -24 when checking data that is not used during the training.

The ability to reconstruct the shape contributes to the results that building shape skeletons are more important in the model than the skeletons of the space between the buildings (Froyen et al., 2015).

The dimensionality problem of linear regression models can be solved by eliminating some of the x variables or by gradually selecting one by one, if it improves the model results. These are stepwise selection models, named forward selection and backward elimination. Such models are often criticized and are recommended to be replaced by newer and more reliable models, such as lasso (Flom, Cassell, 2007). There, the weights are selected for each variable, which allows to exclude even those that improve the model prediction very slightly and discover the synergistic relationships between the variables that improve prediction only in certain combinations. Nevertheless, it could be seen from the results that the coefficient of determination of a test set is negative as well (-6). This shows that with variables, calculated with graph analysis of building and space between the building skeletons, the social capital index cannot be predicted with a linear function.

3. CONCLUSIONS AND POSSIBLE FUTURE DEVELOPMENT

1. There is relation between the urban space as memory and social capital. The relationship is proven with two separate models (LGBM, XGBoost) by using data from separate locations (USA, Lithuania). Machine learning model, using urban space as memory analysis data, was able explain 30% variation of social capital in the USA and Lithuania.

2. The straight skeleton analysis can be used as a model of comprehension of the urban form. This is confirmed by the relationship found in this study between the social capital and the data collected from buildings by skeletal analysis. This is confirmed as well by the analysis of the literature, which shows the versatility and wide application of the skeleton of the form.

3. The existence of the relationship between the urban shape as memory and social capital could start a new understanding of urbanism, where urban forms will be created, considering the factors that influence social capital index. In order to achieve this, the modelling accuracy should be improved, and the influencing factors should be investigated further.

4. The demography of Lithuanian social capital has been an unexplored space so far. The demographic data matrix of Lithuanian social capital that was compiled during the research has already allowed making some insights. In addition, in the future, such comparisons can be made with the data obtained from different analyses of urban space and used for other studies as well.

5. The social capital map shows a long-standing fragile situation where uncontrolled and unplanned urbanization, known as the Urban Sprawl, occurs, which creates massive low-density single-type buildings where the lack of infrastructure burdens the transport systems and reduces the social capital. This is especially relevant in circular municipalities; therefore, the difference in the social capital index between the central and circular municipalities could become a measure of this negative phenomenon.

6. Although the accuracy of the developed models (30%) in the USA and Lithuania is not sufficient for forecasting, it only indicates the significance of the relationship between the urban form and social capital. There are many datasets in the Internet and models of artificial intelligence that are being developed and are fighting for the accuracy of more than 99%; thus, the accuracy of the model that has been developed in this study may seem poor. However, one must be aware that in order to achieve such accuracy, it is necessary to take samples in excess of hundreds of thousands. Moreover, the methodology of this study has specific attachment to the administrative boundaries, which are available in limited numbers, essentially limiting the sample numbers of all datasets. Due to this limitation, it was not possible to create a model with higher accuracy. Nevertheless, such modelling accuracy is common in the social sciences, because large data samples are unavailable and there are many factors that make the data impossible (or too expensive) to collect; further on, it is affected by the complexity of human social behaviour.

7. It is possible to distinguish the features of urban-shaped skeletons that most influenced the formation of the model. These are variations in shape in skeletal depths, branch corners and orientation to countries around the world. Different depths of the shaped skeletons could be obtained from buildings of different planned complexity with few or many enclosures. Different angles in the shaped skeletons are obtained when parts of a building or buildings are oriented towards different angles and when buildings have non-parallel external walls. When visually comparing urban structures, it is possible to select the examples of better and worse social capital, which are visually recognisable for having certain distributions of proportions. Namely, the multiple variations of proportions in a city had high social capital index, and monotonic urban structure had negative; the data analysis and modelling cannot confirm this intuition. Unfortunately, the model cannot explain how these characteristics affect the social capital, i.e., positively or negatively.

8. The properties of skeletons in the form of the buildings had a significantly greater influence on the prediction of the social capital index than the spaces between the buildings. This suggests that the social capital is formed more inside the buildings than in the public spaces. One of the phenomena that is probably related to this observation is climate. This hypothesis could be tested by examining more countries with different climatic conditions, but then, the results may be distorted as well by the influence of different cultures and political situations.

9. One of the available improvements of the model is the possibility of modelling social capital factors separately. Some of the social capital factors are directly related to the purpose of the buildings; thus, the model may be better at predicting these factors. This could be confirmed by the fact that the data collected from the skeletal analysis of buildings predict better than from the spaces between the buildings. In this case, the data that is collected from the spaces between buildings could potentially predict the factors, unrelated to the purpose of the buildings, better.

10. It is likely that in the future, it will be possible to compile a map of social capital, independent of administrative units. As the availability of data in the public and commercial sectors increases, it will be possible to break down the demographics of each factor separately. The study of social capital could be supplemented with data from the virtual social networks. Such fragmentation would allow a map, in which the subdivision would reflect areas of similar social capital index, and smaller areas would allow for a larger sample and most likely create a more accurate model for predicting social capital index from the urban skeleton.

4. SANTRAUKA

4.1. Įžanga

Aktualumas ir motyvacija

Tvarią urbanizaciją sudaro keturios pagrindinės sritys: aplinkosauginė, ekonominė, kultūrinė ir socialinė. Šioms sritims yra skiriamas nevienodas dėmesys ir atrodo, kad socialinė sritis yra labiausiai apleista. Dažnai susidaro įspūdis, kad socialiniai reiškiniai matuojami tik neigiamais rodikliais: nusikalstamumu, socialinio būsto poreikiu, išmokėtų išmokų dydžiu ir pan. Vertinant iš kompleksinių sistemų požiūrio taško, galima pagalvoti, kad vyraujantis požiūris į socialinę sistemą yra tik kaip į kompleksinę sistemą, esančią ekstremaliame, arba kriziniame, režime. Tačiau galbūt ši sistema nėra tokia, tiesiog nėra sukurta pakankamai jautrių rodiklių, matuojančių pozityvias socialinės aplinkos savybes, arba jie nenaudojami. Todėl šio tyrimo tikslas yra atkreipti dėmesį į socialinę miestų aplinką, pritaikyti pasaulyje žinomus rodiklius ir ieškoti priemonių, kaip šiuos socialinius rodiklius paveikti planuojant miestą. Šiuo tikslu aptariami socialiniai rodikliai, akcentuojant vieną svarbiausių – socialinį kapitalą. Yra dvi socialinio kapitalo rūšys: asmeninis ir vietinis. Šis tyrimas yra sutelktas į vietinį socialinį kapitalą, kuris apibūdina bendras vertybes, socialines normas ir pasitikėjimą, o tai lemia visuomeniškumą ir bendradarbiavimą. Asmeninis socialinis kapitalas yra asmeninių ryšių visuma, stiprumas arba galimybė pasinaudoti jais. Nors asmeninį kapitalą dažnai galima traktuoti kaip neigiamą dalyką, vietos socialinis kapitalas yra pozityvesnis. Pirmasis socialinio kapitalo termino pavartojimas ir buvo skirtas vietos socialiniam kapitalui apibūdinti, kai tam tikros socialinės iniciatyvos tampa įmanomos dėl visuotinių asmeninių ryšių pagausėjimo.

Kaip dalis aplinkos determinizmo teorijų socialiniam kapitalui iš dalies galėjo daryti įtaką fizinės miesto erdvės savybės. Perspektyviausia teorija yra traktuoti miestą kaip socialinės sąmonės ar atminties tęsinį. Ji turi teorinius pagrindus elgesio bei psichologijos, taip pat neuromokslų srityse, kurias papildomai sustiprina duomenų mokslas ir vis didėjantis skaitmeninių technologijų pajėgumas. Miesto aplinkos, arba kitaip urbanistinės firmos, traktavimą kaip atminties talpyklą galima būtų kildinti iš teorijos formos kaip atminties talpyklos, kuri bando modeliuoti, kaip žmogus suvokia formas. Nors jau yra daugybė teorijų, leidžiančių modeliuoti reiškinius, kurie vienu metu vyksta miestuose, pavyzdžiui: erdvės sintaksė, izovistas, ląsteliniai automatai, GIS, deja, į miestų planavimą šie metodai įtraukiami retai. Priešingai, miesto planavimo procesas yra inertiškas ir statiškas, o jo metu socialinio kapitalo indeksas nematuojamas, jo nuspėjimui pagal fizinę aplinką išvis nėra jokių įrankių. Nors pastaruoju metu duomenų mokslo ir nuotolinio stebėjimo pažanga lėmė „sumanaus miesto“ sąvokos atsiradimą ir su tuo susijusius pokyčius, kurie pamažu nukreipia miesto planavimą nuo praeityje liekančio sąstingio, tačiau yra tik keli šios filosofijos dalinio įgyvendinimo pavyzdžiai ir nėra išmaniojo miesto pavyzdžio, kuris galėtų pasigirti visišku filosofijos įgyvendinimu. Nors ir yra sukurtų metodų, skirtų įvertinti įvairius miesto formos parametrus, jie daugiausia dėmesio skiria transporto sistemoms ar žemės paskirties tipui, tačiau architektūros vertinimo metodai koncentruojasi į vaizdinės informacijos harmoniją, todėl architektūrai skirtų metodų pritaikymas miesto formai matuoti galėtų padėti sukurti geresnius modelius, kurie leistų numatyti socialinio kapitalo indeksą.

Darbo tikslas

Šios disertacijos tikslas yra paaiškinti ryšį ir dėsningumus, pastebėtus tarp urbanistinės formos kaip atminties talpyklos ir socialinio kapitalo, taip pat pasiūlyti metodiką, naudojant formos skeletą, matematinę grafų analizę bei mašininį mokymąsi siekiant įvertinti urbanistinės formos kaip atminties talpyklos poveikį socialiniam kapitalui.

Disertacijos uždaviniai

Tam kad pasiekti disertacijos tikslą, reikia įvykdyti keletą smulkesnių uždavinių:

- Išanalizuoti esamus metodus ir teorijas, kurios padeda paremti „urbanistinės formos kaip atminties“ sampratą.
- Išanalizuoti Leytono siūlomą metodiką „forma kaip atmintis“ ir pritaikyti ją urbanistikai.
- Ištirti, kaip didžiuosius duomenis galima panaudoti rengiant urbanistinės formos kaip atminties talpyklos metodiką.
- Pritaikyti metodiką, kad ji atitiktų šiuolaikinius dirbtinio intelekto, mašininio mokymosi modelius.
- Įvertinti kelis mašininio mokymosi modelius naudojant užtikrintumo koeficientą ir naudojantis juo išsirinkti tinkamiausią socialinio kapitalo indeksui prognozuoti. Tai atlikti panaudojant laisvai prieinamus socialinio kapitalo indekso duomenis JAV teritorijai.
- Apskaičiuoti urbanistinės formos kaip atminties talpyklos parametrus, bei patikrinti, ar socialinio kapitalo modeliavimas yra įmanomas antroje pasirinktoje vietoje (Lietuvoje).
- Paaiškinti urbanistinės formos kaip atminties talpyklos ryšio su socialiniu kapitalu dėsningumus.

Ginamieji teiginiai

- Urbanistinė forma kaip atminties talpykla yra statistiškai reikšmingai susijusi su socialiniu kapitalu.
- Teorinę „urbanistinės formos kaip atminties talpyklos“ sampratą galima praktiškai išmatuoti naudojant sukurtą tiesiųjų formos skeletų analizės metodiką, kuri gali analizuoti pastatų bei erdvių tarp pastatų formas.
- Žmogaus erdvės suvokimą galima apytiksliai aproksimuoti su formos skeleto modeliu.
- „Socialinio kapitalo“ matavimo metodika administraciniuose vienetuose pakankamai universali, kad galėtų būti pritaikoma

skirtingoms valstybėms, ir pateikia duomenis, kurie savaime yra naudingi suprantant urbanistines tendencijas.

Mokslinis naujumas

Nors yra daug tyrimų, teigiančių, kad fizinė aplinka turi įtakos socialiniams rodikliams, tačiau tie tyrimai koncentravosi tik į labai siaurą socialinį aspektą arba buvo nedidelės apimties. Tyrimų, kurie kūrė modelius, galinčius iš fizinės miesto aplinkos prognozuoti socialinį kapitalą, rasti nepavyko. Taip pat nebuvo rasta ir jokių tyrimų, kurie miesto formą vertintų kaip atminties talpyklą. Šis tyrimas pristato originalią metodiką, skirtą urbanistikos tyrimams bei naudojančią formos skeleto matematinę grafų analizę ir mašininį mokymąsi. Formos skeletas sukuriamas metodu, vadinamu skeletonizacija, kurio metu yra randamos centrinės arba ašinės linijos, kurių visuma atspindi supaprastintą topologinę formos struktūrą. Ši metodika anksčiau nebuvo naudojama analizuojant urbanistinę formą ir prognozuojant socialinę statistiką, pavyzdžiui, socialinį kapitalą

Praktinis panaudojimas

Teorija „urbanistinė forma kaip atminties talpykla“ leidžia naujai įvertinti urbanistiką. Nors yra kelios metodikos, skirtos miesto formos vertinimui, ši išsiskiria savo sąsaja su fizinės aplinkos supratimo įtaka žmonių elgesiui. Urbanistinė forma kaip atminties talpykla galėtų būti naudojama ne tik socialiniam kapitalui modeliuoti, bet ir kitiems svarbiems ir nebūtinai socialiniams rodikliams, susijusiems su urbanistika, pavyzdžiui: pėsčiųjų eismui, nusikalstamumo numatymui, laiko tėkmės suvokimui, vietos parinkimo uždaviniams ir daugeliui kitų. Modelis, leidžiantis prognozuoti socialinį kapitalą išskaičiuojant pirminius duomenis tik iš fizinės miesto aplinkos, galėtų padėti sukurti precedentų neturinčią urbanistinio planavimo metodologiją, kuri leistų miesto planą optimizuoti socialiniam kapitalui gerinti. Galima išskirti urbanistinės formos kaip atminties talpyklos faktorius, kurie labiausiai veikia socialinį kapitalą, ir paaiškinti jų poveikį. Toks faktorių sąrašas su paaiškinimu įgalintų žmogiškos intuicijos atsiradimą ir įžvalgų, rastų naudojant modelį, taikymą be paties modelio.

Nors yra sukurta daugybė modelių, siūlančių įvairias miesto formos nulemtas prognozes, nė vienas iš jų nėra bendrųjų planų procese, todėl ši metodika galėtų būti pridėta prie kritinės miesto metodikos masės. Tikėtina, kad pasiekus kritinę moksliškų metodikų urbanistiniam planavimui masę, bendrųjų planų procesas nebegalės jų ignoruoti ir jie bus įtraukti į proceso eigą ir pakels jį į naują aukštesnį ir sumanesnį lygį.

Metodologija

Disertacija tyrimo eigoje yra naudojami keli metodai. Literatūros analizė, naudojant indukcines ir dedukcines išvadas. Jomis remiantis pasirinkti duomenų surinkimo, analizės bei modeliavimo metodai. Urbanistinė forma kaip atminties talpykla yra naujas transdisciplinis urbanistikos modelis, suvienijantis savyje psichologijos, matematinės grafų analizės, neuronų skaičiavimo, statistinio modeliavimo ir architektūros teorijos mokslo šakas. Socialinio kapitalo indeksas matuojamas pagal sociologijos srities metodiką, kurios pagrindas yra statistinės analizės metodai, įskaitant pagrindinio komponento analizę. Socialinio kapitalo ir urbanistinės formos kaip atminties talpyklos duomenys yra erdviniai, todėl juos reikia suvienyti naudojant geografines informacines sistemas (GIS). Galutinis rezultatas, modelio užtikrintumo koeficientas matuojamas išbandant įvairius mašininio mokymosi modelius, tokius kaip linijinė regresija, LASSO, dirbtiniai neuroniniai tinklai, gilusis mokymasis, LTSM ir XGBoost. Duomenų analizės modelių tikslumas įvertinamas kryžminės validacijos metu užtikrintumo koeficientu R^2. Visa metodika yra glaudžiai susieta, atskirus tyrime naudojamus metodus būtų galima priskirti multidisciplininiams, todėl naudojant kartu jie tyrimą padaro transdisciplininiu.

4.2. Mokslinių tyrimų apžvalga bei analizė

Ryšio tarp urbanistinės formos kaip atminties talpyklos ir socialinio kapitalo esmė gali slypėti fundamentiniuose moksluose, net ir tuose, kurie anksčiau nebuvo naudojami urbanistikoje. Tyrimo metu nuorodų apie tokio ryšio egzistavimą buvo ieškoma ne tik srityse, tiesiogiai susijusiose su urbanistika, bet ir susijusiose su supančios erdvės suvokimu ir jo skeliamais jausmais, nes jausmai gali nulemti veiksmus, o visų įmanomų veiksmų aibėje yra ir, šiam tyrimui patys svarbiausi, socialiniai veiksmai.

Kad būtų galima sukurti urbanistinės formos kaip atminties talpyklos analizavimo bei socialinio kapitalo nuspėjimo modelius, reikia mokėti juos išmatuoti arba aprašyti skaičiais ar parametrais. Tam tikslui pasiekti apžvelgiami istoriniai bandymai išmatuoti ar modeliuoti aplinką ir jos savybes, susijusias su aplinkos suvokimu, kuris yra svarbus urbanistinei formai kaip atminties talpyklai. Nors estetika ir navigacija iš pirmo žvilgsnio neturi nieko bendro tarpusavyje, tačiau žvelgiant iš urbanistinės formos kaip atminties talpyklos pozicijos, šios sferos yra susijusios aplinkos suvokimo aspektu. Todėl analizuojant įvairias sritis, susijusias su pažinimo aplinka, būtų galima įgyti supratimą, kaip ją panaudoti tyrimuose, nepaisant to, kad atskirose studijose daugiausia dėmesio skiriama estetikai ar navigacijai. Terminas „miesto forma kaip atmintis“ yra kilęs iš Leytono termino „forma kaip atmintis“, kuriame teigiama, kad kiekviena forma turi savo sukūrimo istoriją, kurią teoriškai būtų galima išgauti analitiniu būdu. Anot šios teorijos, žmogaus smegenys naudoja šį procesą suvokdamos ir įsiminimos aplinką. Nors ši teorija yra įdomi, jos praktiniam panaudojimui trukdo faktas, kad nėra sukurta modelio, kuris projektuoti šį procesą. Erdvės pažinimą galima traktuoti trimis būdais: tiesiog filosofiškai, nesistengiant rasti jokių pritaikymų, bandant atspėti taisykles, kurias gamta ar žmogus naudojo kurdami aplinką, ir bandyti rasti matavimo metodus, turinčius sąsajų su aplinkos suvokimu. Kadangi suvokimas gali būti vertinamas kaip išplaukiantis biologinių smegenų neuronų visuomenės rezultatas, miestas gali būti vertinamas kaip organizmas, o žmonių bendruomenė formuoja nuoseklų socialinio kapitalo rezultatą. Todėl galima rasti įrodymų neurobiologijoje arba eksperimentų su gyvūnais rezultatuose. Taip pat galima ištirti dirbtinės aplinkos kūrimo kultūras ir tai, kaip buvo imtasi tam tikros strategijos kuriant aplinką, kuri įkvepia socialinę sąveiką.

„Loci metodas“ geriau žinomas kaip atminties rūmų vaizduotės technika, kurią naudojant įsivaizduojama aplinka, kurioje saugomi atsiminimai (Yates, 2014). Tai, kad naudojant šį metodą pagerėja gebėjimas įsiminti, yra paprasčiausias ir geriausias įrodymas, kad aplinkos suvokimas yra fundamentalus žmogaus atminties procesas.

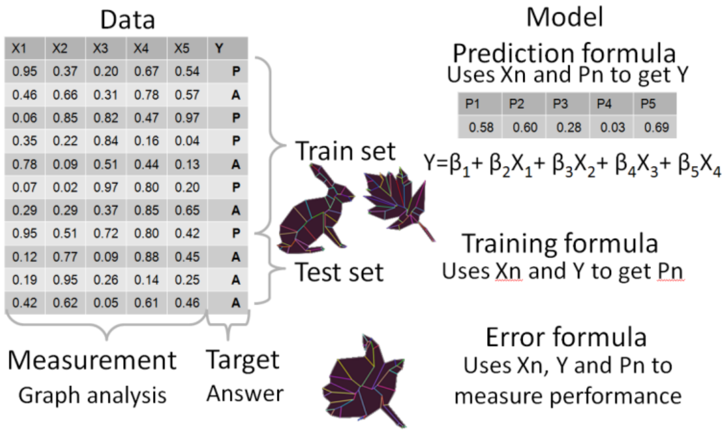
Terminas „urbanistinė erdvė kaip atminties talpykla“ savaime nurodo į miesto erdvės būseną, kuri egzistuoja tik atmintyje, tokiame stovyje, kai ji suvokimo procesas vykstantis žmogaus smegenyse ją apdorojo į formą, kurią atmintyje saugo žmogaus smegenys. Yra tyrimų, analizuojančių miesto erdvę iš šios perspektyvos, žymiausias „Miestovaizdis“ (Lynch, 1960). Iš esmės ši teorija teigia, kad žmonių įsimintas miestas skiriasi nuo tikrovės, bet taip pat ir priklauso ir nuo žmogaus, vis dėlto kai kurie pasikartojantys elementai išlieka. Teorija koncentruojasi į kelis aspektus: kelius, ribas, rajonus, mazgus ir orientyrus. Keliai yra atpažįstamos jungtys (gatvės, takai, kanalai), kuriomis asmuo galėtų judėti. Ribos yra atpažįstamos sienos, dalijančios kelias zonas. Rajonai yra tos zonos. Mazgai yra centrinės vietos, kurios dažniausiai yra pageidaujamos susitikimo ar susirinkimo vietos. Jie paprastai, bet nebūtinai turi orientyrus, kurie tampa jų tapatybės dalimi. Orientyrai, išskirtiniai gamtos ar žmogaus sukurti objektai (skulptūros, paminklai ir kt.) nebūtinai turi būti mazguose. Kai kurie šių elementų deriniai gali turėti sinerginį poveikį, kiti gali trukdyti vienas kitam. Elementų trūkumas sukuria tapatumo spragą, kuri iškreipia miesto psichinį vaizdą. Nors ši teorija turi savo originalią metodiką, susidedančią iš kokybinių socialinių tyrimų, ji buvo modernizuota ir pritaikyta skaitmeniniam amžiui tokiais metodais kaip izovistas (Benedikt, 2016). Panašių teorijų yra ir daugiau, todėl jų požiūrio į erdvės supratimo mechanizmą analizė galėtų suteikti geresnį supratimą apie urbanistinę formą kaip atminties talpyklą ir jos poveikį socialiam kapitalui.

Grožio ir harmonijos jausmas tikriausiai yra įgimtas. Evoliucinė psichologija teigia, kad jis išsivystė, nes buvo naudingas mūsų priešistorinių protėvių išgyvenimui (Dutton, 2009), tačiau, tai labai sunku paaiškinti, išmatuoti ar kiekybiškai įvertinti. Nuo seniausių laikų mąstytojai bandė suprasti šį jausmą ir paaiškinti sąlygas, kurios jį sukėlė. Idealus tokio bandymo rezultatas būtų analitinis metodas, leidžiantis iš nesunkiai pamatuojamų pradinių duomenų prognozuoti estetizmo pojūtį reprezentuojančią tolydžiąją vertę. Paprasčiausias būdas tai pasiekti yra atspėti ar sugalvoti taisyklę, kuri galėtų paaiškinti procesą ar numatyti jausmo stiprumą. Nors taisyklėmis paremtų modelių rezultatai dažnai yra gruboki, tačiau jie yra lengviausiai suvokiami ir paaiškinami, todėl gali suteikti pradinių žinių ir gali nurodyti teisingą kelią ieškant optimaliausio estetikos suvokimo modelio, kuris siejasi su urbanistinės formos kaip atimties talpyklos teorija.

4.3. Metodologija

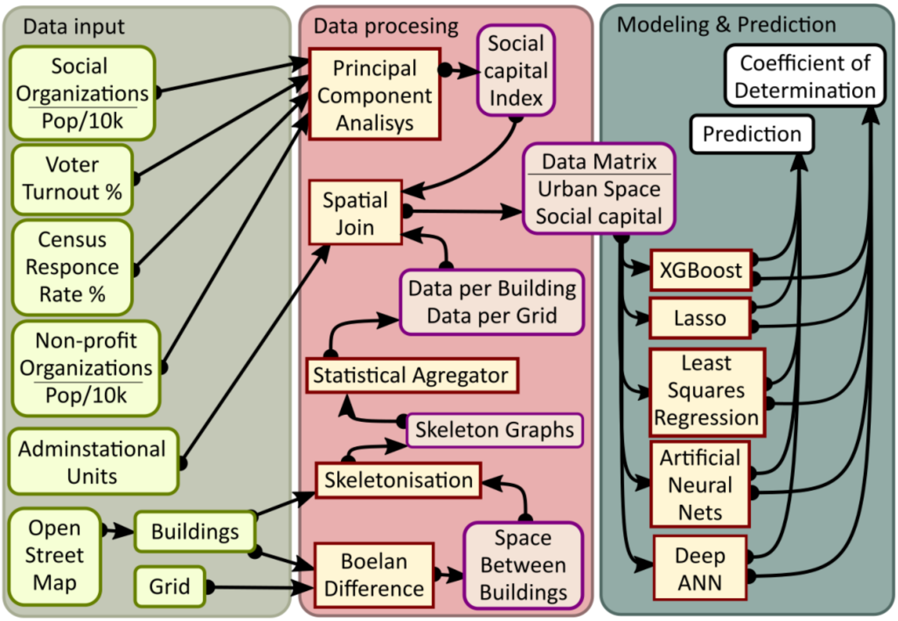
Galutinis šio tyrimo tikslas yra rasti ryšį tarp urbanistinės formos kaip atminties talpyklos ir socialinio kapitalo ir paaiškinti, kokie faktoriai ir kaip jį įtakoja. Urbanisinė forma, kaip atminties talpykla, turi keletą komponentų, todėl prieš pradedant kurti pagrindinį modelį reikia papildomai išbandyti jų universalumą. Yra trys pagrindiniai metodai, kuriuos galima pasirinkti matuojant urbanistinę formą kaip atmintį: adaptyvių griaučių naudojimas (Wilder, Feldman, & Singh, 2011a), tiesaus skeleto naudojimas arba adaptyvių griaučių aproksimavimas. Adaptyvūs griaučiai gali būti per sunkiai įgyvendinami nuo nulio, nes autoriai nepateikia kodo, be to, teigia, kad jis yra labai lėtas, todėl netinka masiniam skeletonizavimui, reikalingam šiems tyrimams. Tiesusis skeletas yra tinkamiausias miesto formoms, nes jos paprastai yra tiesios, taip pat yra patikimas įgyvendinimas, priešingai, tai labai neįprastoje aplinkoje, kurioje yra „PostgreSQL“ duomenų bazė (PostGIS, 2018a).

Daugumai mašininio mokymosi modelių reikia, kad duomenys būtų parengti duomenų matricoje. Modeliai, kuriems netaikomi tokie reikalavimai, paprastai specializuojasi tam tikro tipo duomenims, kuriuos neefektyvu laikyti duomenų matricoje, kai kurie pavyzdžiai yra vaizdai, tekstas bei garso įrašai.



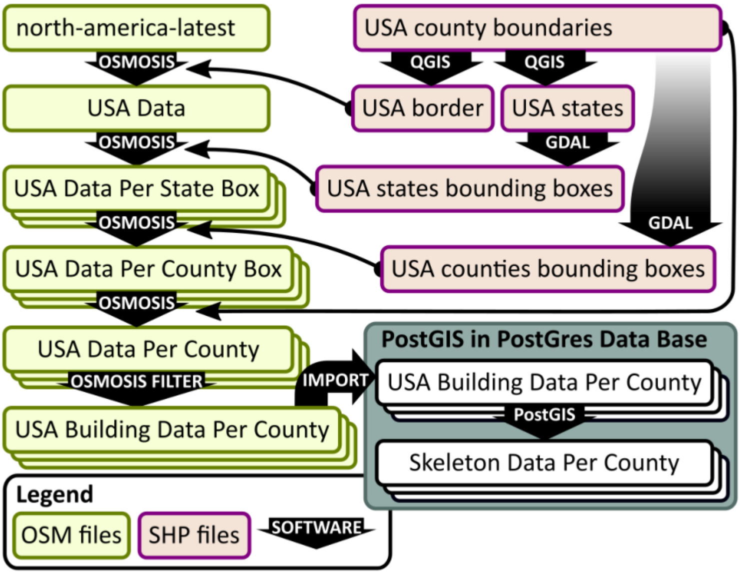
13 pav. Mašininio apsimokymo pavyzdys (tiesinė regresija)

Putnamo socialinio kapitalo indekso skaičiavimo metodika yra pagrįsta statistikos metodais ir savaime sukuria duomenų matricą, nors skalė yra skirtinga, ji sukuria vieną indeksą kiekvienam administraciniam vienetui, pavyzdžiui, apskričiai, o naudojant pažangiausias skeleto technologijas, neįmanoma gauti šios didelės teritorijos griaučių. Panagrinėkime situaciją, kad rastume šio apribojimo sprendimą. Norint išanalizuoti miesto formą kaip atmintį, galima sukurti tiesų skeletą kiekvienam pastatui, tuo tarpu erdvėms tarp pastatų reikalinga kita strategija. Buvo pasirinkta padalinti plotą į paprastą kvadratinį tinklelį, kurį būtų galima panaudoti iškerpant pastatą. Gautas tinklelis su pastatų pjūviu galėtų atstoti erdvę tarp pastatų, be to, jiems būtų galima sukurti tiesiuosius skeletus. Remiantis literatūros analize buvo pasirinktas 300 m tinklelio dydis.



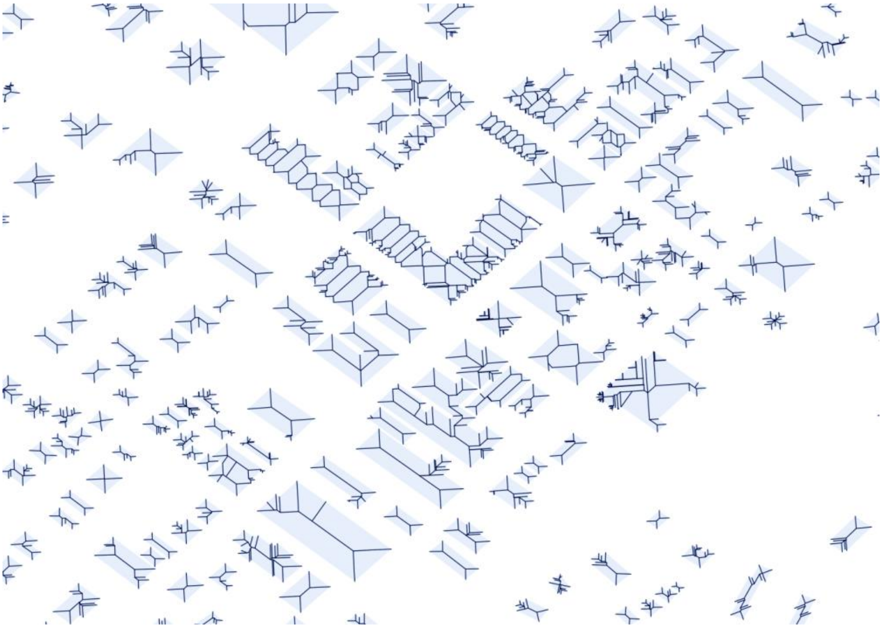
14 pav. Metodologijos apžvalgine diagrama

Pastatų ir erdvių tarp pastatų tiesiesiems skeletams sugeneruoti buvo panaudoti duomenys iš savanorių kuriamo atviro GIS žemėlapio „OpenStreetMap“ (“OpenStreetMap”, 2019).



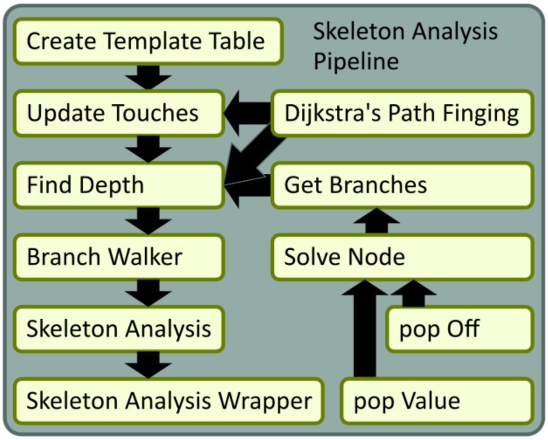
15 pav. Duomenų paruošimo eigos diagrama

Duomenų filtravimas paliekant tik pastatų tipą ir supjaustymas į atskiras apskritis buvo atliekamas naudojant atvirojo kodo programinės įrangos įrankį (Osmosis, 2016). Sprendžiant iš Feldmano pateikto aprašymo, šioje programoje jis būtų skaičiavimo požiūriu per brangus (Feldman & Singh, 2006). Todėl skeleto skaičiavimui buvo pasirinktas „PostGIS“ plėtinys SFCGAL (SFCGAL, 2016; PostGIS, 2018b). SFCGAL skeleto įgyvendinimas yra pagrįstas (Aichholzer, Aurenhammer, Alberts, & Gärtner, 1996). Programa Osmosis turi prisijungimo į „PostGIS“ duomenų bazę funkciją, kuria buvo pasinaudota įkeliant pastatų poligonus. Skeletams apskaičiuoti buvo naudojama SFCGAL funkcija „ST_StraightSkeleton“ sistemoje „PostGIS“. Tokiu būdu buvo sukurta duomenų bazės struktūra, kur kiekvienam administraciniam vienetui buvo sukurta nauja lentelė, kurioje kiekviena eilutė atitiko pastatą ir geometrijos stulpelyje buvo įrašyta tiesiojo skeleto reprezentacija (15 pav.).



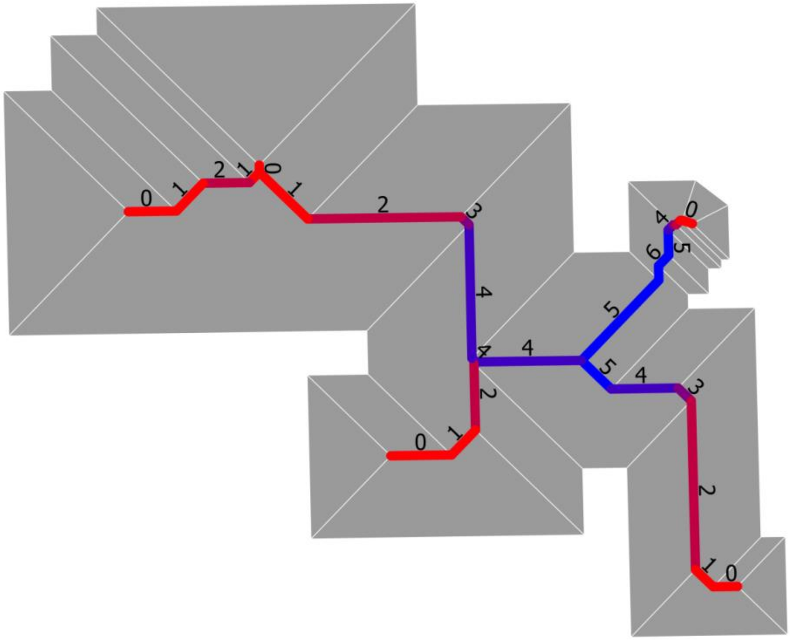
16 pav. Pastatų tiesieji skeletai

Norint išanalizuoti informaciją apie tiesiųjų skeletų mazgus, saugomą vienos langelio duomenų bazės įraše viduje, buvo sukurta speciali „PostGIS“ funkcija, kuri sukurdavo duomenų bazėje laikinę lentelę, kurioje eilutė reprezentuodavo tiesiojo skeleto kaip matematinio grafo mazgą. Analizavimo procesas paremtas Wilderio aprašymu užpildydavo laikiną lentelę tarpinėmis analitinėmis reikšmėmis analizavimo proceso metu (Wilder, Feldman, & Singh, 2011b).



17 pav. Tiesiojo skeleto analizės eiga

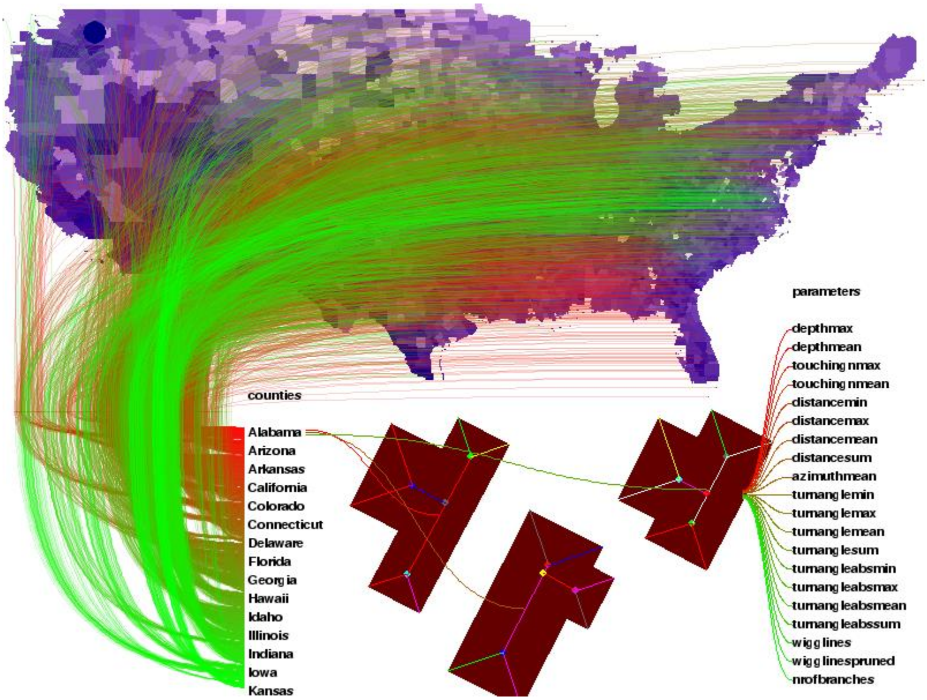
Wilderio aprašytas analitinis procesas yra paremtas įsivaizduojama agento, judančio tiesiojo skeleto kaip matematinio grafo ryšiais nuo vieno mazgo iki kito pradedant nuo giliausio mazgo. Giliausias mazgas buvo randamas pasinaudojant „Dijkstra“ algoritmo kelio paieškos algoritmo variantas, kuriame visiems matematinio grafo mazgams buvo suteikta gylio reikšmė pradedant tuo mazgu su viena jungtimi, kuriam buvo priskirtas nulinis gylis (Dijkstra, 1959). Algoritmas keliauja jungtimis į grafos gylį ir užpildo gylio reikšmę praeitų jungčių kiekiu, o susidūrus dviem šakoms pasirenkama mažesnė gylio reikšmė. Taip surandamas giliausias matematinio grafo mazgas, kuris yra pirmas analitinis parametras, bet taip pat atskaitos taškas, nuo kurio atliekamos kitos analizės, kuriose randamas išsišakojimų kiekis, posūkių skaičius ir kampai. Šios mazgų reikšmės yra agreguojamos statistinėmis funkcijomis: minimumu, maksimumu ir vidurkiu, kad būtų gauta papildomų galutinės analizės rodiklių. Papildomai skaičiuojamas sukauptas posūkių kampas bei sukauptas absoliutus posūkio kampas sudedant posūkių kampus agento kelionėje nuo matematinio grafos centro iki kraštų, iš kurių paskaičiuojamas grafo susisukimo parametras dalinant posūkių kampų sumą iš absoliučių posūkių kampų sumos.



26 pav. Apgenėto tiesiojo skeleto (raudonai-mėlyno) spalvos ir skaičiai reprezentuoja briaunos gylį

Wilderio naudojamas skeletų generavimo algoritmas veikia optimizavimo principu, panašiai kaip anksčiau aprašytas regresijos modelis, ir turi metaparametrus, kurie leidžia sukonfigūruoti algoritmą optimaliausio skeleto radimui, tačiau šiame tyrime naudojamas SFCGAL algoritmas neturi tokių savybių. Norint kompensuoti šį nesutapimą buvo naudojamas Bai pasiūlytas metodas (Bai, Latecki ir Liu, 2007). Jo esmė yra minimalaus gylio skeleto mazgų ištrynimas arba „apgenėjimas“ ir analitinio algoritmo pakartojamas analizuojant naujai gautą matematinį grafą. Nugenėti tiesiųjų skeletų matematiniai grafai turi gilesnę topologiją ir neturi šakų, nereprezentuojančių ašinių linijų (17 pav.).

Tokiu būdu gauti analitiniai duomenys yra kiekvienam pastatui, o socialinio kapitalo duomenys yra kiekvienam administraciniam vienetui, kuris gali turėti įvairų pastatų skaičių. Tokių duomenų neįmanoma sujungti į viena duomenų matricą. Norint sujungti šiuos duomenis reikia agreguoti pastatų skeletų analizės rezultatus į administracinius vienetus, paskaičiuojant tokius statistinius parametrus kaip minimumas, maksimumas bei vidurkis. Tokiu būdu gauta duomenų matrica turėjo 116 stulpelių.

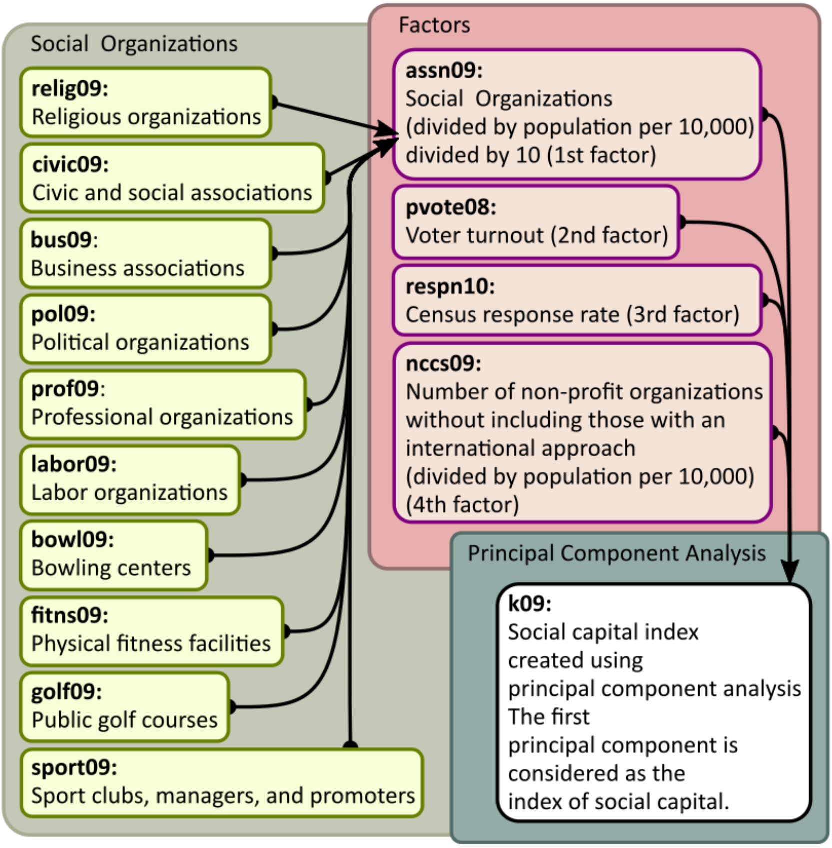


27 pav. Parametrų agregavimo schema: tiesaus skeleto analizės rezultatai (apačioje kairėje) agreguoti ir priskirti administraciniams vienetams (viršuje)

Duomenų analizė buvo atlikta pasinaudojant atvirojo kodo paketu „Scikit-learn“ (Cournapeau, 2018a). Dauguma pastatų buvo paprastų formų, stačiakampiai gretasieniai, tačiau didėjant pastatų formų sudėtingumui jų skaičius mažėjo. Salingaros ir Westas teigia, kad aplinka, kuri neatitinka universalaus mastelio ir pasiskirstymo, neigiamai veikia sveikatą (Salingaros & West, 1999). Universalus mastelis ir pasiskystamas gali būti sulyginamas su kitomis teorijomis, iš kurių turbūt geriausiai žinoma yra laipsnių dėsnis (angl.: *power law*) (Clauset, Shalizi, & Newman, 2009). Šis dėsnis ir yra skirtas analizuoti situacijas, kuriose mažų reikšmių yra labai daug ir reikšmių skaičius mažėja didėjant pačioms reikšmėms. Tam kad būtų patikrinta, ar laipsnių dėsnis turi įtakos ir socialiniam kapitalui, naudojantis atviro kodo įrankiu duomenų matrica buvo papildyta laipsnio dėsnio statistikomis (Alstott, Bullmore, & Plenz, 2013). Įrankis sugeneruoja keturias statistikas, todėl panaudojus ji gautą duomenų bazę sudarė 2400 mėginių, gavusių 880 parametrų.

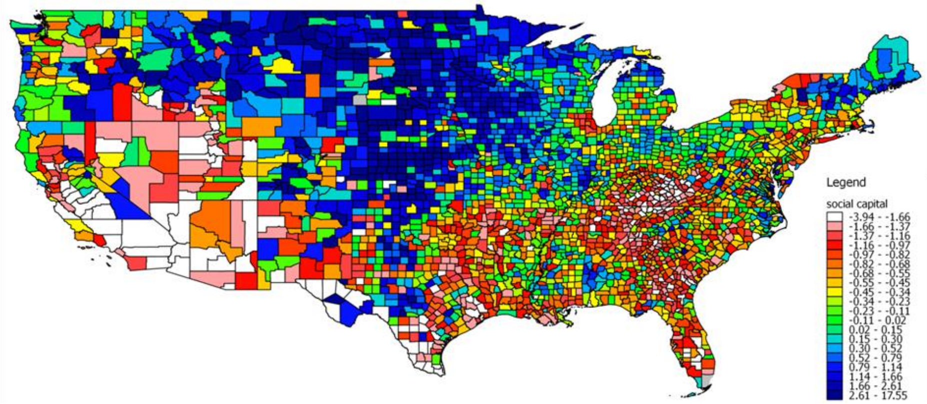
Vienas iš būdų statistiškai įrodyti ryšio tarp urbanistinės formos kaip atminties talpyklos ir socialinio kapitalo yra modelio, gebančio modeliuoti socialinio kapitalo indeksą naudojant tik analitinius urbanistinės formos kaip atminties talpyklos duomenis, sukūrimas. Naudojantis skirtumu tarp tikrojo ir prognozuojamo socialinio kapitalo galima apskaičiuoti užtikrintumo rodiklį, kuris savaime reprezentuoja modelio tikslumą bei ryšio stiprumą, bet taip pat leidžia lyginti modelius tarpusavyje (Colin Cameron & Windmeijer, 1997). Paprasčiausias ir geriausiai žinomas modelis – mažiausių kvadratų regresija – buvo naudojama kaip pradinis atspirties taškas, kurio rezultatas buvo bandomas gretinti su kitais modeliais (Geladi & Kowalski, 1986). Taip pat buvo naudojamas universaliu aproksimatoriumi laikomas daugiasluoksnis perceptronas (Rumelhart, Hinton, & Williams, 1985). Geriausius rezultatus parodė jungtinio gradiento atsitiktinių medžių svoriniai kolektyviniai modeliai. Rezultatai buvo papildomai pagerinti. Galima dar labiau padidinti prognozes pašalinant kai kuriuos parametrus taikant laipsninio parametru pasirinkimo algoritmą (Koller & Sahami, 1996).

Socialinio kapitalo duomenims buvo naudojama atviros prieigos duomenų bazė (Rupasingha, Goetz, & Freshwater, 2006).



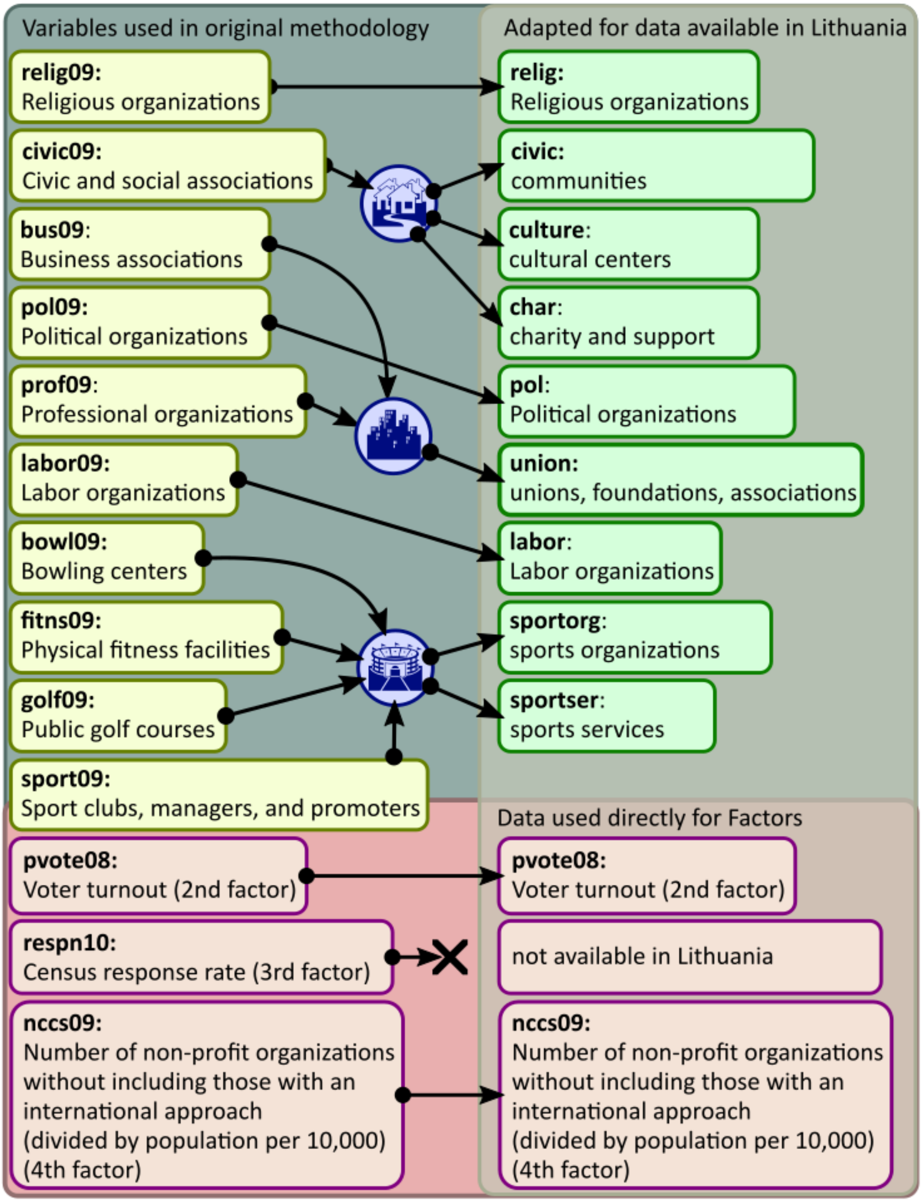
28 pav. Socialinio kapitalo indekso nustatymo metodologija (Rupasingha et al., 2006)

Šie duomenys buvo sukurti vadovaujantis Putnamo teorija (Putnam, 2001). Duomenų bazę sudaro įvesties kintamieji ir galutinė apskaičiuota socialinio kapitalo indekso vertė. Ji apima JAV valstijas, išskyrus Aliaską. Joje yra įrašas kiekvienai apskričiai kiekvienoje valstybėje, iš viso 3 108 įrašai.



29 pav. JAV Socialinio kapitalo indeksas. Duomenys iš (Rupasingha et al., 2006)

Pasirinktais metodais Lietuvoje nebuvo matuotas nei socialinis kapitalas, nei urbanistinė forma. Šiam tyrimui atlikti reikėjo sudaryti specialią duomenų matricą, kuri sujungtų abiejų rūpimų komponentų GIS duomenis. Socialinio kapitalo indeksas apskaičiuojamas naudojant pagrindinės komponentės metodą su socialinių struktūrų, nepelno organizacijų, visuotinio surašymo aktyvumo ir balsavimo aktyvumo faktoriais. Tam, kad būtų galima rasti socialinių struktūrų faktorių, reikia suskaičiuoti, kiek kiekviename administraciniame vienete yra: religinių organizacijų, socialinių bei pilietinių asociacijų, verslo asociacijų, profesinių organizacijų, darbininkų organizacijų, boulingo centrų, viešų sporto aikštelių, viešų golfo aikštelių, sporto klubų. Šių struktūrų sumą reikia padalinti iš 10 ir to administracinio vieneto populiacijos 10 000 dalies. Tam, kad būtų galima rasti nepelno organizacijų faktorių, reikia suskaičiuoti, kiek yra tokių organizacijų ir padalinti iš to administracinio vieneto populiacijos 10 000 dalies. Balsavimo aktyvumas yra tiesiog santykis balsavusiųjų ir visų turinčių teisę balsuoti kiekviename administraciniame vienete. Visuotinio surašymo aktyvumas yra santykis atsakiusiųjų į surašymo anketą su administracinio vieneto populiacija.



30 pav. Socialinio kapitalo metodologijos pritaikymas Lietuvai

Pats sudėtingiausias atkartojimo prasme yra pirmasis faktorius, į kurį įeina įvairių organizacijų skaičius. Sunku suvokti, kodėl boulingo centrai, viešos sporto aikštelės, viešos golfo aikštelės ir sporto klubai yra išskirti į atskiras kategorijas. Galima būtų manyti, kad siekta pabrėžti jų svarbą, tačiau matematiškai toks išskyrimas nieko nekeičia, nes vėliau kintamųjų reikšmių suma dalinama iš kintamųjų skaičiaus. Galima būtų teigti, jog apskaičiuojamas vidurkis. Galbūt kintamųjų atrinkimui šiam faktoriui apskaičiuoti įtakos turėjo ir duomenų prieinamumas. Detaliausią klasifikaciją turinčioje Lietuvos organizacijų duomenų bazėje yra 286 klasės (Rekvizitai.lt, 2019). Artimiausios Putnam metodui atrinktos šios: religinės organizacijos; bendruomenės; kultūros centrai; labdara ir parama; sąjungos, fondai, asociacijos; socialinės paslaugos; sporto organizacijos; sporto paslaugos; politinės organizacijos. Tam, kad būtų galima gauti faktoriaus reikšmes, šių organizacijų kiekis buvo apskaičiuotas kiekvienai savivaldybei ir padalintas iš kiekvienos savivaldybės populiacijos 10 000 dalies. Nepelno organizacijų faktoriui buvo panaudotos VŠĮ įmonės iš „info.lt" katalogo (HNIT-Baltic Duomenys, 2015). Analogiškai nepelno organizacijos buvo suskaičiuotos kiekvienai savivaldybei ir padalintos iš kiekvienos savivaldybės populiacijos 10 000 dalies. Pagal Lietuvos statistikos departamento visuotinio surašymo metodiką, būtina pasiekti 94 % populiacijos. Deja, į anketą atsako mažiau ir likę duomenys yra užsakomi iš VĮ Registrų centro. Pagal tą pačią metodiką duomenys apie tai, kokia dalis atsakė, yra nekaupiami (stat.gov.lt, 2011). Dėl to Lietuvos socialinio kapitalo nustatymui šio faktoriaus negalime naudoti, duomenų jam nėra. Tačiau tai nėra problema, nes principinę komponentę galima išvesti iš bet kokio faktorių skaičiaus. Dėl to socialinio kapitalo tikslumas tikriausiai sumažės arba nebus toks pilnavertis. Nepaisant to, jis bus apskaičiuotas pasirinktu prisitaikančiu principinių komponenčių metodu. Balsavimo aktyvumo faktoriui buvo panaudoti Lietuvos Respublikos vyriausiosios rinkimų komisijos teikiami duomenys apie 2016 m. Seimo rinkimus (vrk.lt, 2016). Duomenys yra pateikiami balsavimo apygardomis, kurių ribos nevisiškai sutampa su savivaldybių ribomis, duomenys savivaldybėms buvo perkelti naudojantis apygardų geometriniais centrais. Taip buvo sumuojamas apygardų, kurių geometriniai centrai įeina į savivaldybę skaičius, kuris padalinamas iš populiacijos, kad būtų gauta balsavusiųjų proporcija. Duomenys buvo apjungti naudojantis „Qgis" geografinio informacinio paketo „Join attributes by location" funkcija („QGIS", 2017). Pagrindinė komponentė buvo apskaičiuota naudojantis „Python" programavimo kalbos automatino apsimokymo paketu „sklearn" (Rossum & Drake, 2011; Cournapeau, 2018b).

Kartu sudėjus pastatų ir erdvių tarp pastatų kintamuosius, buvo sudaryta duomenų matrica iš 1750 kintamųjų ir 60 imčių, nes Lietuvoje yra 60 savivaldybių. Socialinio kapitalo modeliavimui vien tik su Lietuvos duomenimis buvo sudaryti įvairūs modeliai: tiesinė regresija, „lasso", jungtinio gradiento atsitiktinių medžių svoriniai kolektyviniai modeliai „lgbm" ir „xgboost" (Microsoft Corporation, 2019; Chen, 2016). Tam, kad būtų galima patikrinti modelio tikslumą, buvo atskirta atsitiktinė 20 % duomenų dalis, kuri apsimokinimo metu modeliui buvo nerodoma, tačiau tikslumas buvo skaičiuojamas tik pagal mokymo metu nuslėptą dalį.

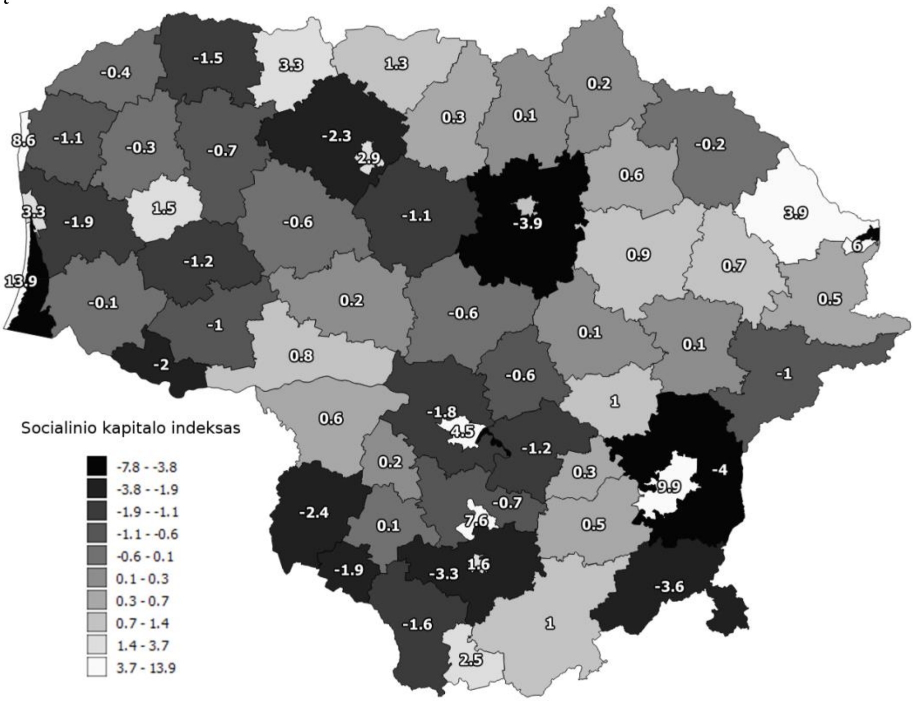
Jau buvo minėta, kad imčių kiekis sutampa su administracinių teritorijų kiekiu (60), tačiau formos skeletų analizės rezultatas yra 1750 kintamųjų. Toks mažas imčių kiekis labai apriboja automatinio apsimokymo modelio pasirinkimą. Taip pat faktas, kad kintamųjų yra daugiau nei imčių, reiškia, kad modelis turės tendenciją prisitaikyti prie duomenų (persimokyti), kitaip tai žinoma kaip dimensiškumo problema, kuri labai apriboja jo galimybę prognozuoti (Verikas & Gelžinis, 2003). Tą parodo ir didelis skirtumas tarp modelio testavimo su duomenimis, naudotais apmokymui, ir naudotais tik testavimui. Pavyzdžiui, „xgboost“ modelio patikimumo koeficientas su apmokymo duomenimis yra 0,8, tačiau su testavimo duomenimis tik 0,3. Tai reiškia, kad didžiąja dalimi modelis yra persimokęs. To buvo galima tikėtis, nes kaip parodė modeliavimo patikimumo ryšio su duomenų kiekiu tyrimai, šiam modeliui apsimokyti nepakanka imčių (Floares et al., 2017). Turint daugiau imčių, šiuo atveju modeliuojant daugiau administracinių vienetų, galima būtų manipuliuoti modelio hiperparametrais tam, kad priartėtų apsimokymo ir testavimo patikimumo rodikliai.

1 lentelė. Modelių tikslumo palyginimas

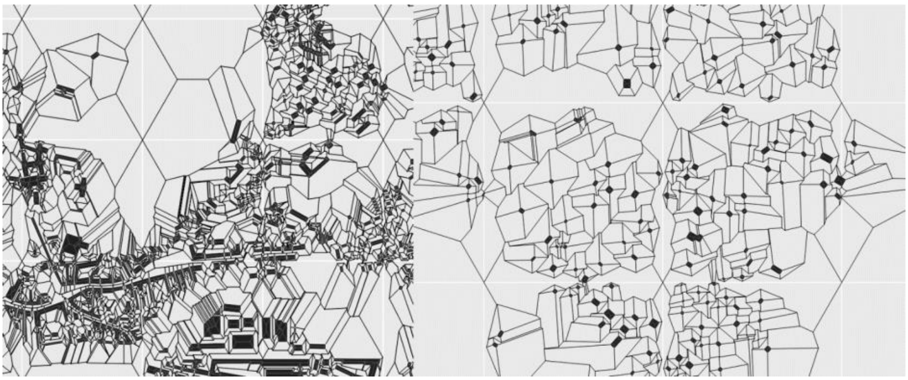
Model	Coeficient of determination (R^2)
Linear Regresion	-24
lasso	-6
lgbm	0,1
xgboost	0,3

Atsitiktinių medžių svoriniai kolektyviniai modeliai, naudojami su hipoteze, kad matematinėmis funkcijomis duomenų matricos modeliuoti neįmanoma. Buvo išbandyti tiesinės regresijos ir „lasso“ modeliai, kurių labai mažas patikimumo koeficientas bent jau dalinai patvirtino šia hipotezę. Atsitiktinių medžių modelis kiekvieną „x“ tolydųjį kintamąjį bando paversti į diskretųjį, rasdamas ribą (tolydžiojo kintamojo reikšmę), kurią naudoja tolydžiojo kintamojo suklasifikavimui. Jei toks naujas diskretusis kintamasis padeda prognozavimui, jis paliekamas ir ieškoma kita ribos reikšmė, kuri įves papildomų kategorijų diskrečiajame kintamajame ir, tikėtina, dar labiau pagerins prognozavimo galimybes. Tai yra atsitiktinio medžio gylis. Toks pats procesas pakartojamas kiekvienam kintamajam. Kadangi toks modelis bando kategorizuoti ir tikrina prognozavimo pagerėjimą kiekvienam kintamajam atskirai, jis neturi dimensiškumo problemos, taip pat gali įvertinti kiekvieno kintamojo klasifikavimo naudingumą. Šiame tyrime buvo naudoti 4 žingsnių gylio atsitiktiniai medžiai. Svoriniai kolektyviniai modeliai gali jungti vieno tipo ar netgi keleto skirtingų tipų modelių prognozes įvertindami jų sklaidas ir pagal jas sumažindami arba padidindami atskiro modelio prognozės įtaką galinei, suminei prognozei. Tai leidžia iš daug silpnai prognozuojančių modelių sudaryti vieną stipriau prognozuojantį. Šiame tyrime geriausias rezultatas buvo pasiektas naudojant 10 000 narių svorinį kolektyvą.

Interpretuodami Lietuvos socialinio kapitalo žemėlapį, galime pastebėti kelias ir taip žinomas tendencijas. Žiedinėse savivaldybėse, miestų rajonuose yra ekstensyvios plėtros kvartalų. Buvę žemės ūkio paskirties sklypai sudalinti į smulkius namų valdos sklypelius. Tokiems kvartalams ar net rajonams trūksta infrastruktūros ir jie skatina transporto sistemos apkrovą. Žiedinės ir centrinės savivaldybių socialinių kapitalų skirtumas galėtų būti situacijos aštrumo matas. Priešinga tendencija yra su kurortais, kuriuose yra daugiau infrastruktūros nei jos reikia vietiniams gyventojams, tai irgi prisideda prie transporto sistemos apkrovimo. Tačiau būtų klaidinga teigti, kad tokių teritorijų socialinio kapitalo indeksas yra neteisingas, jis yra tiesiog epizodinis. Birštono savivaldybė turi pastebimą abiejų paminėtų tendencijų sinergiją, ji yra ir kurortas, ir centrinė savivaldybė, turinti savo žiedinę.



31 pav. Lietuvos socialinio kapitalo indekso žemėlapis



32 pav. Vilniaus miesto skeletai (kairėje; soc. kapitalo indeksas 9,9) ir Vilniaus rajono (dešinėje; soc. kapitalo indeksas -4)

Stebint skirtingo socialinio kapitalo indekso urbanizuotas teritorijas, galima intuityviai pastebėti tam tikrų urbanistinių morfotipų ir genotipų asociacijas su mažu arba dideliu socialiniu kapitalu. Putnam metodika kartu su prieinamų duomenų specifika pririša socialinio kapitalo matavimą prie administracinių vienetų, tačiau pačiame administraciniame vienete socialinis kapitalas gali būti nevienalytis. Galbūt ateityje bus prieinami smulkesni ir gausesni duomenys, kurie leis išskirti teritorijų su vienalyčiu socialiniu kapitalu ribas bei socialinio kapitalo židinius.

4.4. Išvados ir galima tolimesnė eiga

1. Tarp miesto erdvės kaip atminties ir socialinio kapitalo yra silpnas statistinis ryšys. Ryšys įrodomas dviem atskirais modeliais (LGBM, XGBoost), naudojant duomenis iš atskirų vietų (JAV, Lietuva). Mašinų mokymosi modelis, naudojantis urbanistinę formą kaip atminties talpyklos analizės duomenis, galėjo paaiškinti 30 % socialinio kapitalo variacijos JAV ir Lietuvoje.

2. Formos skeletas gali būti naudojamas kaip urbanistinės formos suvokimo modelis. Tai patvirtina šiame tyrime rastas ryšys tarp socialinio kapitalo ir formos skeleto analize surinktų duomenų iš pastatų. Taip pat tai patvirtina ir literatūros analizė, kurioje parodomas formos skeleto universalumas ir platus pritaikymas.

3. Ryšio tarp urbanistinės formos kaip atminties talpyklos ir socialinio kapitalo egzistavimas leidžia plėtoti naują urbanistikos šaką, kur miesto formos kuriamos taip pat atsižvelgiant ir į socialinio kapitalo indekso veiksnius. Tačiau tam, kad tokia praktika galėtų prasidėti, reikia pagerinti modeliavimo tikslumą ir geriau ištirti veiksnius, turinčius įtakos.

4. Lietuvos socialinio kapitalo demografija iki šiol buvo neištirta. Tyrimo metu sudaryta Lietuvos socialinio kapitalo demografijos duomenų matrica jau leido daryti kai kurias įžvalgas net nesugretinant jos su kitais duomenimis. Be to, ateityje galima atlikti tokius sugretinimus su duomenimis, gautais iš skirtingų urbanistinės erdvės analizių, taip pat panaudoti kitiems tyrimams.

5. Socialinio kapitalo žemėlapyje įžvelgiama nuo seno žinoma opi situacija, kai nekontroliuojama ir nesuplanuota urbanizacija, žinoma kaip ekstensyvi plėtra, sukuria didžiulius žemo tankio vienos paskirties užstatymus, kuriuose infrastruktūros įvairovės nebuvimas apkrauna transporto sistemas ir mažina socialinį kapitalą. Tai ypač aktualu žiedinėse savivaldybėse, todėl socialinio kapitalo indekso skirtumas tarp centrinės ir žiedinės savivaldybės galėtų tapti šio negatyvaus reiškinio matu.

6. Nors sukurto modelio tikslumas (30 %) nėra pakankamas prognozavimui, jis nurodo ryšio stiprumą tarp urbanistinės formos ir socialinio kapitalo. Šiuo metu, kai interneto platybėse yra daugybė duomenų rinkinių ir kuriami dirbtinio intelekto modeliai kaunasi dėl tikslumo viršijančio 99 %, šio tyrimo sukurto modelio tikslumas gali atrodyti menkas. Vis dėlto reikia turėti omenyje, kad norint pasiekti tokį tikslumą reikia turėti imtis, viršijančias šimtus tūkstančių įrašų, o dėl šio tyrimo metodikos specifinio prisirišimo prie administracinių ribų ir kartu su jų ribotu skaičiumi imtis taip pat labai ribota. Dėl šio apribojimo nebuvo galima sukurti modelio su didesniu tikslumu. Nepaisant to, socialiniuose moksluose toks modeliavimo tikslumas yra įprastas ne tik dėl to, kad didelės duomenų imtys yra neprieinamos, bet ir dėl daugybės faktorių, kurių duomenų neįmanoma (ar per brangu) surinkti, tačiau jie įtakoja žmonių socialinį elgesį.

7. Galima išskirti urbanistinės formos skeletų savybes, kurios labiausiai įtakojo modelio sudarymą. Tai yra variacijos formos skeletų gyliuose, šakų kampuose ir orientacijoje su pasaulio šalimis. Skirtingi formos skeletų gyliai galėtų būti gaunami iš pastatų, kurie yra skirtingo planinio sudėtingumo, turintys mažai arba daug korpusų. Skirtingi kampai formos skeletuose gaunami, kai pastato dalys arba korpusai yra orientuoti skirtingais kampais vienas kito atžvilgiu, taip pat, kai pastatai turi nelygiagrečias išorines sienas. Deja, modelis negali paaiškinti, kaip šios savybės veikia socialinį kapitalą: teigiamai ar neigiamai.

8. Pastatų formos skeletų savybės turėjo žymiai didesnę įtaką socialinio kapitalo indekso prognozavimui nei erdvių tarp pastatų. Vadinasi, socialinis kapitalas yra labiau formuojamas pastatų viduje, o ne viešosiose erdvėse. Tai gali būti ir lokalus reiškinys, susijęs su Lietuvos klimatu. Šią hipotezę galima būtų bandyti patikrinti ištiriant daugiau valstybių su skirtingomis klimatinėmis sąlygomis, tačiau tada rezultatus gali iškreipti ir skirtingų kultūrų bei politinių situacijų įtaka.

9. Tobulinant modelį, galima bandyti sumodeliuoti socialinio kapitalo faktorius atskirai. Dalis socialinio kapitalo faktorių yra tiesiogiai susiję su pastatų paskirtimi, dėl to gali būti, kad modeliui geriau sekasi prognozuoti šiuos faktorius. Tai patvirtintų ir tai, kad duomenys, surinkti iš pastatų skeletų analizės prognozuoja geriau nei iš erdvių tarp pastatų. Tokiu atveju duomenys, surinkti iš erdvių tarp pastatų, potencialiai galėtų geriau prognozuoti faktorius, nesusijusius su pasatų paskirtimi.

10. Tikėtina, kad ateityje galima bus sudaryti socialinio kapitalo žemėlapį, nepriklausomą nuo administracinių vienetų. Didėjant duomenų prieinamumui valstybiniuose ir komerciniuose sektoriuose, galima bus smulkinti kiekvieno faktoriaus demografiją atskirai. Taip pat socialinio kapitalo tyrimas galėtų būti papildytas virtualių socialinių tinklų duomenimis. Toks susmulkinimas leistų nubraižyti žemėlapį, kuriame teritorijų suskirstymas atspindėtų panašaus socialinio kapitalo indekso teritorijas, taip pat smulkesnės teritorijos leistų turėti didesnę imtį ir, tikėtina, sukurti tikslesnį socialinio kapitalo indeksą prognozuojantį modelį iš urbanistinės formos skeleto.

5. REFERENCES

1. 3ds Max History. (2011, May 4). . Retrieved December 18, 2018, from https://web.archive.org/web/20110504184716/http://wiki.cgsociety.org/index.php/3ds_Max_History

2. Abelson, H., Goodman, N., & Rudolph, L. (1974). LOGO Manual. Retrieved August 8, 2019, from https://dspace.mit.edu/handle/1721.1/6226

3. Adamatzky, A. (1996). Voronoi-like partition of lattice in cellular automata. Retrieved February 17, 2020, from https://www.sciencedirect.com/science/article/pii/0895717796000039

4. Adler, P. S., & Kwon, S.-W. (2002). Social Capital: Prospects for a New Concept. The Academy of Management Review, 27(1), 17–40.

5. Adorno, T. W. (1998). Aesthetic Theory. (R. Hullot-Kentor, Ed.) (1 edition.). Univ Of Minnesota Press.

6. Aichholzer, O., Aurenhammer, F., Alberts, D., & Gärtner, B. (1996). A Novel Type of Skeleton for Polygons.

7. Akiba, F. (2013). Preface: Natural Computing and Computational Aesthetics. In Y. Suzuki & T. Nakagaki (Eds.), Natural Computing and Beyond, Proceedings in Information and Communications Technology (pp. 117–118). Tokyo: Springer Japan.

8. Alexander, C. (1977). A Pattern Language: Towns, Buildings, Construction. New York: Oxford University Press.

9. Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., & Angel, S. (1977). A Pattern Language: Towns, Buildings, Construction. New York: Oxford University Press.

10. Almeida, C. M., Gleriani, J. M., Castejon, E. F., & Soares-Filho, B. S. (2008). Using neural networks and cellular automata for modelling intra-urban land-use dynamics. International Journal of Geographical Information Science, 22(9), 943–963. Taylor & Francis.

11. Alstott, J., Bullmore, E., & Plenz, D. (2013). Powerlaw: A Python package for analysis of heavy-tailed distributions. ArXiv:1305.0215 [physics]. Retrieved August 7, 2019, from http://arxiv.org/abs/1305.0215

12. Anderson, W. P., Kanaroglou, P. S., & Miller, E. J. (1996). Urban Form, Energy and the Environment: A Review of Issues, Evidence and Policy. Urban Studies, 33(1), 7–35. SAGE Publications Ltd.

13. Appert-Rolland, C., Degond, P., & Motsch, S. (2011). Two-way multi-lane traffic model for pedestrians in corridors. Networks and Heterogeneous Media, 6(3), 351–381.

14. Archives: Keynote Speech to the 1996 OOPSLA Convention. (1996). . Retrieved June 23, 2018, from http://www.patternlanguage.com/archive/ieee.html

15. Bach, M., Meigen, T., & Strasburger, H. (1997). Raster-scan cathode-ray tubes for vision research-limits of resolution in space, time and intensity, and some solutions. Spatial Vision, 10(4), 403–414.

16. Bai, X., Latecki, L. J., & Liu, W.-Y. (2007). Skeleton pruning by contour partitioning with discrete curve evolution. IEEE transactions on pattern analysis and machine intelligence, 29(3), 449–462.

17. Barnes, J. A. (1969). GRAPH THEORY AND SOCIAL NETWORKS: A TECHNICAL COMMENT ON CONNECTEDNESS AND CONNECTIVITY. Sociology, 3(2), 215–232.

18. Barnett, J. (2011). A short guide to 60 of the newest urbanisms and there could be many more, 77, 19–21.

19. Barnsley, M. F., & Sloan, A. D. (1990, July 10). Methods and apparatus for image compression by iterated function system. Retrieved August 14, 2019, from https://patents.google.com/patent/US4941193/en

20. Batty, M., Xie, Y., & Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. Computers, Environment and Urban Systems, 23(3), 205–233.

21. Batty, Michael, & Longley, P. (1994a). Fractal Cities: A Geometry of Form and Function (1 edition.). London: Academic Press.

22. Batty, Michael, & Longley, P. (1994b). Fractal Cities: A Geometry of Form and Function (1 edition.). London: Academic Press.

23. Bays, C. (2012). Cellular Automata in Triangular, Pentagonal and Hexagonal Tessellations. In R. A. Meyers (Ed.), Computational Complexity: Theory, Techniques, and Applications (pp. 434–442). New York, NY: Springer New York. Retrieved July 28, 2019, from https://doi.org/10.1007/978-1-4614-1800-9_28

24. Beck, K., & Cunningham, W. (1987). Using Pattern Languages for Object-Oriented Programs. Retrieved June 23, 2018, from http://c2.com/doc/oopsla87.html

25. Beggs, J. M., & Plenz, D. (2003). Neuronal Avalanches in Neocortical Circuits. Journal of Neuroscience, 23(35), 11167–11177.

26. Beier, S., Bernard, J.-B., & Castet, E. (2018). Numeral legibility and visual complexity. Presented at the DRS Design Research Society, 2018: Limerick. Retrieved July 22, 2019, from https://adk.elsevierpure.com/da/publications/numeral-legibility-and-visual-complexity

27. Benedikt, M. L. (2016). To Take Hold of Space: Isovists and Isovist Fields: Environment and Planning B: Planning and Design. Retrieved February 12, 2020, from https://journals.sagepub.com/doi/10.1068/b060047

28. Biedl, T., Held, M., Huber, S., Kaaser, D., & Palfrader, P. (2015). Weighted straight skeletons in the plane. Computational Geometry, 48(2), 120–133.

29. Birkhoff, G. D. (1933). Aesthetic Measure. Retrieved February 2, 2020, from https://www.hup.harvard.edu/catalog.php?isbn=9780674734470

30. Black, L. (2014). Does John Conway Hate Life? Retrieved August 15, 2019, from https://www.i-programmer.info/news/112-theory/7029-does-john-conway-hate-life.html

31. Black, P. (1998). Dictionary of Algorithms and Data Structures. Retrieved June 23, 2018, from https://xlinux.nist.gov/dads/

32. Blatnig, S. (2009). Microscopic Traffic Simulation with Intelligent Agents: Simulation of Human Driving Behaviour. Saarbrücken, Germany: VDM Verlag Dr. Müller.

33. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3(Jan), 993–1022.

34. Blum, H. (1964). A Transformation for Extracting New Descriptors of Shape. Retrieved July 25, 2019, from http://isgwww.cs.uni-magdeburg.de/~stefans/npr/entry-Blum-1967-ATF.html

35. Bourdieu, P. (1986). The forms of capital. In: Richardson, J., Handbook of Theory and Research for the Sociology of Education. Westport, CT: Greenwood: 241–58. Retrieved August 5, 2019, from http://www.socialcapitalgateway.org/content/paper/bourdieu-p-1986-forms-capital-richardson-j-handbook-theory-and-research-sociology-educ

36. Broadbent, N. J., Squire, L. R., & Clark, R. E. (2006). Reversible hippocampal lesions disrupt water maze performance during both recent and remote memory tests. Learning & Memory, 13(2), 187–191.

37. Cacciola, F. (2004). A CGAL implementation of the Straight Skeleton of a Simple 2D Polygon with Holes. ResearchGate. Retrieved July 25, 2019, from https://www.researchgate.net/publication/240939888_A_CGAL_implementation_of_the_Straight_Skeleton_of_a_Simple_2D_Polygon_with_Holes

38. Carpenter, F., Manson, D., Jeffery, K., Burgess, N., & Barry, C. (2015). Grid cells form a global representation of connected environments. Current biology: CB, 25(9), 1176–1182.

39. Chen, J.-S. (1995). Fractal image compression based on visual perception. Human Vision, Visual Processing, and Digital Display VI (Vol. 2411, pp. 92–99). Presented at the Human Vision, Visual Processing, and Digital Display VI, International Society for Optics and Photonics. Retrieved August 14, 2019, from https://www.spiedigitallibrary.org/conference-proceedings-of-spie/2411/0000/Fractal-image-compression-based-on-visual-perception/10.1117/12.207530.short

40. Chen, T. (2016). Story and Lessons Behind the Evolution of XGBoost. Retrieved May 31, 2019, from https://homes.cs.washington.edu/~tqchen/2016/03/10/story-and-lessons-behind-the-evolution-of-xgboost.html

41. Cheng, S.-W., & Vigneron, A. (2002). Motorcycle Graphs and Straight Skeletons. Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA ’02 (pp. 156–165). Philadelphia, PA, USA: Society for Industrial and Applied Mathematics. Retrieved July 25, 2019, from http://dl.acm.org/citation.cfm?id=545381.545401

42. Ching, F. D. K. (2007). Architecture: Form, Space, and Order (3rd edition.). Hoboken, N.J: John Wiley & Sons.

43. Chomsky, N. (1956). Three models for the description of language—IEEE Journals & Magazine. Retrieved August 12, 2019, from https://ieeexplore.ieee.org/document/1056813

44. Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. SIAM Review, 51(4), 661–703.

45. Cole, M. (1996). Cultural psychology: A once and future discipline. Cultural psychology: A once and future discipline. Cambridge, MA, US: Harvard University Press.

46. Colin Cameron, A., & Windmeijer, F. A. G. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. Journal of Econometrics, 77(2), 329–342.

47. Corbusier, C.-É. (1923). Toward An Architecture. Retrieved August 19, 2019, from https://www.abebooks.com/9780892368990/Architecture-Corbusier-Cohen-Jean-louis-introduction-0892368993/plp

48. Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). Introduction to Algorithms. MIT Press.

49. Cornea, N. D., Silver, D., Yuan, X., & Balasubramanian, R. (2005). Computing hierarchical curve-skeletons of 3D objects. The Visual Computer, 21(11), 945–955.

50. Cournapeau, D. (2018a). Scikit-learn: Machine learning in Python—Scikit-learn 0.21.2 documentation. Retrieved May 31, 2019, from https://scikit-learn.org/stable/

51. Cournapeau, D. (2018b). Scikit-learn: Machine learning in Python—Scikit-learn 0.21.2 documentation. Retrieved May 31, 2019, from https://scikit-learn.org/stable/

52. Cudeiro, J., & Sillito, A. M. (2006). Looking back: Corticothalamic feedback and early visual processing. Trends in Neurosciences, 29(6), 298–306.

53. Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2006). Studying Aesthetics in Photographic Images Using a Computational Approach. In A. Leonardis, H. Bischof, & A. Pinz (Eds.), Computer Vision – ECCV 2006, Lecture Notes in Computer Science (pp. 288–301). Berlin, Heidelberg: Springer.

54. Datta, S., & Mukhopadhyay, S. (2013). A Composite Method Based on Formal Grammar and DNA Structural Features in Detecting Human Polymerase II Promoter Region. PLoS ONE, 8(2). Retrieved August 12, 2019, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3577817/

55. Davies, T. P., Fry, H. M., Wilson, A. G., & Bishop, S. R. (2013). A mathematical model of the London riots and their policing. Scientific Reports, 3, 1303.

56. Davis, W. (1992). Masking the Blow: The Scene of Representation in Late Prehistoric Egyptian Art (y First edition edition.). Berkeley: University of California Press.

57. Dechter, R. (1986). Learning while searching in constraint-satisfaction-problems. Proceedings of the Fifth AAAI National Conference on

Artificial Intelligence, AAAI’86 (pp. 178–183). Philadelphia, Pennsylvania: AAAI Press.

58. Demir, I., Hahn, C., Leonard, K., Morin, G., Rahbani, D., Panotopoulou, A., Fondevilla, A., et al. (2019). SkelNetOn 2019: Dataset and Challenge on Deep Learning for Geometric Shape Understanding. ArXiv:1903.09233 [cs]. Retrieved July 22, 2019, from http://arxiv.org/abs/1903.09233

59. Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. Numerische Mathematik, 1(1), 269–271.

60. Dixon, R. (1986). Aesthetics: A Cognitive Account. Leonardo, 19(3), 237–240.

61. Doeller, C. F., Barry, C., & Burgess, N. (2010). Evidence for grid cells in a human memory network. Nature, 463(7281), 657–661.

62. Dutton, D. (2005). Aesthetics and Evolutionary Psychology. The Oxford Handbook of Aesthetics. Retrieved February 12, 2020, from https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199279456.001.0001/oxfordhb-9780199279456-e-41

63. Dutton, D. (2009). Aesthetics and Evolutionary Psychology. ResearchGate. Retrieved February 3, 2020, from https://www.researchgate.net/publication/292481220_Aesthetics_and_Evolutionary_Psychology

64. Dye, D. S. (1974). Chinese Lattice Designs (2nd edition.). New York: Dover Publications.

65. Egli, S., & Crossland, D. (2013). Metapolator – Free Font Editor for Designing Typeface Families. Retrieved July 23, 2019, from http://metapolator.com/home

66. Eilouti, B. (2019). Shape grammars as a reverse engineering method for the morphogenesis of architectural façade design. Frontiers of Architectural Research, 8(2), 191–200.

67. Euclid. (300AD). The Elements. (T. L. Heath, Tran.) (1956th ed.). USA: Dover Publications, Inc.

68. Euler, L. (1741). Solutio problematis ad geometriam situs pertinentis. Commentarii academiae scientiarum Petropolitanae, 128–140.

69. Everitt, B., & Howell, D. (2005). Encyclopedia of Statistics in Behavioral Science | Reference Reviews | Vol 20, No 4. Retrieved December 7, 2018, from https://www.emeraldinsight.com/doi/abs/10.1108/09504120610664123

70. Falconer, K. (2004). Fractal Geometry: Mathematical Foundations and Applications. John Wiley & Sons.

71. Faria, J., Bagley, S., Rüger, S., & Breckon, T. (2013). Challenges of finding aesthetically pleasing images. 2013 14th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS) (pp. 1–4). Presented at the 2013 14th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS).

72. Feldman, J., & Singh, M. (2006). Bayesian estimation of the shape skeleton. Retrieved July 25, 2019, from https://www.pnas.org/content/103/47/18014

73. Feynman, R. P. (1949). Space-Time Approach to Quantum Electrodynamics. Physical Review, 76(6), 769–789.

74. Figueiredo, L., & Amorim, L. E. C. (2005). Continuity lines in the axial system.

75. Firestone, C., & Scholl, B. J. (2014). "Please tap the shape, anywhere you like": Shape skeletons in human vision revealed by an exceedingly simple measure. Psychological Science, 25(2), 377–386.

76. Flemming, U. (2016). The Secret of the Casa Giuliani Frigerio. Environment and Planning B: Planning and Design. Retrieved February 11, 2020, from https://journals.sagepub.com/doi/10.1068/b080087

77. Flom, P., & Cassell, D. (2007). Stopping Stepwise: Why Stepwise and Similar Selection Methods Are Bad, and What You Should Use. Retrieved from https://www.lexjansen.com/cgi-bin/xsl_transform.php?x=nesug2007

78. Friedman, N., Ito, S., Brinkman, B. A. W., Shimono, M., DeVille, R. E. L., Dahmen, K. A., Beggs, J. M., et al. (2012). Universal Critical Dynamics in High Resolution Neuronal Avalanche Data. Physical Review Letters, 108(20), 208102.

79. Froyen, V., Feldman, J., & Singh, M. (2015). Bayesian Hierarchical Grouping: Perceptual grouping as mixture estimation. Psychological review, 122(4), 575–597.

80. Galton, F. (1907). Vox Populi. Nature. News, . Retrieved August 19, 2019, from https://www.nature.com/articles/075450a0

81. Gardner, M. (1970). Mathematical Games.The fantastic combinations of John Conway's new solitaire game "life." Scientific American. Retrieved August 15, 2019, from http://red.gnoss.com/en/community/vidaartificial/resource/mathematical-gamesthe-fantastic-combinations-of/25dd1ff9-c827-44d6-ad49-da98b22c31bc

82. Gareau, F. H. (1987). Expansion and Increasing Diversification of the Universe of Social Science. International Social Science Journal, 39(4), 595–606.

83. Garrison, W. L., & Marble, D. F. (1962). The Structure of Transportation Networks. ResearchGate. Retrieved August 18, 2020, from https://www.researchgate.net/publication/235082695_The_Structure_of_Transportation_Networks

84. Gehl, J., & Rogers, L. R. (2010). Cities for People (1st Edition edition.). Washington, DC: Island Press.

85. Geladi, P., & Kowalski, B. R. (1986). Partial least-squares regression: A tutorial. Analytica Chimica Acta, 185, 1–17.

86. Geofabrik. (2016). . Retrieved August 6, 2019, from http://download.geofabrik.de/north-america.html

87. GeoPandas. (2018). GeoPandas is an open source project to make working with geospatial data in python easier. Retrieved July 12, 2019, from http://geopandas.org/

88. Goldberg, R., & Romano, F. J. (2000). Digital Typography: Practical Advice for Getting theType You Want When You Want It (1 edition.). San Diego, CA: Windsor Professional Information.

89. Gore, A., & Gupta, S. (2015). Full reference image quality metrics for JPEG compressed images. AEU - International Journal of Electronics and Communications, 69(2), 604–608.

90. Grave, E., Bojanowski, P., Gupta, P., Joulin, A., & Mikolov, T. (2018). Learning Word Vectors for 157 Languages. ArXiv:1802.06893 [cs]. Retrieved August 5, 2019, from http://arxiv.org/abs/1802.06893

91. Gray, H. (1858). Anatomy: Descriptive and surgical (1st ed.). London: J.W. Parker.

92. Gunji, Y.-P., Nishiyama, Y., & Adamatzky, A. (2012). Robust Soldier Crab Ball Gate. ArXiv:1204.1749 [cs, nlin]. Retrieved May 31, 2019, from http://arxiv.org/abs/1204.1749

93. Hagerhall, C. M., Laike, T., Taylor, R. P., Küller, M., Küller, R., & Martin, T. P. (2008). Investigations of Human EEG Response to Viewing Fractal Patterns. Perception, 37(10), 1488–1494.

94. Halatsch, J., Kunze, A., & Schmitt, G. (2008). Using Shape Grammars for Master Planning. In J. S. Gero & A. K. Goel (Eds.), Design Computing and Cognition ’08 (pp. 655–673). Dordrecht: Springer Netherlands.

95. Hanifan, L. J. (1916). The Rural School Community Center. Retrieved June 1, 2019, from https://journals.sagepub.com/doi/10.1177/000271621606700118

96. Hao, K. (2019). We analyzed 16,625 papers to figure out where AI is headed next. MIT Technology Review. Retrieved July 21, 2020, from https://www.technologyreview.com/2019/01/25/1436/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next/

97. Hardy, G. H. (Godfrey H. (1921). A course of pure mathematics. Cambridge, At the University Press. Retrieved February 1, 2020, from http://archive.org/details/coursepuremath00hardrich

98. Hartbauer, M., & Römer, H. (2016). Rhythm Generation and Rhythm Perception in Insects: The Evolution of Synchronous Choruses. Frontiers in Neuroscience, 10. Retrieved August 19, 2019, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4885851/

99. Heaton, J. (2019). Evolving continuous cellular automata for aesthetic objectives. Genetic Programming and Evolvable Machines, 20(1), 93–125.

100. Hegel, G. W. F. (1807). The Phenomenology of Spirit. (T. Pinkard & M. Baur, Trans.). Place of publication not identified: Cambridge University Press.

101. Held, M., & Palfrader, P. (2017). Straight skeletons with additive and multiplicative weights and their application to the algorithmic generation of roofs and terrains. Computer-Aided Design, 92, 33–41.

102. Hendrickx, M., & Wagemans, J. (1999). A Critique of Leyton's Theory of Perception and Cognition. Review of Symmetry, Causality, Mind, by Michael Leyton. Journal of Mathematical Psychology, 43(2), 314–345.

103. Henley. (1993). Statics of a "self-organized" percolation model. Physical review letters, 71(17), 2741–2744.

104. Heudin, J.-C. (1996). A New Candidate Rule for the Game of Two-dimensional Life. Retrieved August 18, 2019, from https://www.complex-systems.com/abstracts/v10_i05_a03/

105. Hillier, B. (1989). The Social Logic of Space (Reprint edition.). Cambridge: Cambridge University Press.

106. Hillier, B., Burdett, R., Peponis, J., & Penn, A. (1987). Creating Life: Or, Does Architecture Determine Anything? ResearchGate. Retrieved January 19, 2020, from https://www.researchgate.net/publication/32884797_Creating_Life_Or_Does_Architecture_Determine_Anything

107. Hillier, P. B., & Sahbaz, O. (2008). An evidence based approach to crime and urban design Or, can we have vitality, sustainability and security all at once?, 28.

108. HNIT-Baltic Duomenys. (2015). Įmonės. Retrieved May 31, 2019, from https://www.arcgis.com/home/item.html?id=8cdfd4ebc61b44e29c67372c44986959

109. Holtan, M., Dieterlen, S., & Sullivan, W. (2015). Social Life Under Cover: Tree Canopy and Social Capital in Baltimore, Maryland. Retrieved July 10, 2019, from https://journals.sagepub.com/doi/abs/10.1177/0013916513518064

110. Hoshika, S., Leal, N. A., Kim, M.-J., Kim, M.-S., Karalkar, N. B., Kim, H.-J., Bates, A. M., et al. (2019). Hachimoji DNA and RNA: A genetic system with eight building blocks. Science (New York, N.Y.), 363(6429), 884–887.

111. Huber, S. (2011). Computing Straight Skeletons and Motorcycle Graphs: Theory and Practice, 134.

112. Huber, S., & Held, M. (2011). Computing Straight Skeletons and Motorcycle Graphs: Theory and Practice, 134.

113. Hung, C.-C., Carlson, E. T., & Connor, C. E. (2012). Medial Axis Shape Coding in Macaque Inferotemporal Cortex. Neuron, 74(6), 1099–1113.

114. Huys, Q. J. M., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., Dayan, P., et al. (2015). Interplay of approximate planning strategies. Proceedings of the National Academy of Sciences of the United States of America, 112(10), 3098–3103.

115. Huysman, M. H., & Wulf, V. (2004). Social Capital and IT, current debate and research. Social Capital and IT, 1–17.

116. Inkscape. (2020). . Retrieved July 21, 2020, from https://inkscape.org/about/

117. Ito, M., Izui, M., Yamazaki, Y., & Matsushita, M. (2003). Morphological Diversity in Crystal Growth of L-Ascorbic Acid Dissolved in Methanol. Journal of the Physical Society of Japan, 72(6), 1384–1389.

118. Itoh, K., & Ohno, Y. (1993). A Curve Fitting Algorithm for Character Fonts. ResearchGate. Retrieved August 5, 2019, from https://www.researchgate.net/publication/220180125_A_Curve_Fitting_Algorithm_for_Character_Fonts

119. Jacob, S. (1594). Rechenbuch auff den Linien und mit Ziffern, sampt allerley Vortheilen, fragweise. Retrieved January 30, 2020, from https://www.e-rara.ch/zut/6765407

120. Jiang, L., Ling, D., Zhao, M., Wang, C., Liang, Q., & Liu, K. (2018). Effective Identification of Terrain Positions from Gridded DEM Data Using Multimodal Classification Integration. ISPRS International Journal of Geo-Information, 7(11), 443.

121. Kaniari, A. (2013). D'Arcy Thompson's On Growth and Form and the Concept of Dynamic Form in Postwar Avant-Garde Art Theory. Interdisciplinary Science Reviews, 38(1), 63–73.

122. Kelly, G. K. (2007). Citygen: An Interactive System for Procedural City Generation.

123. Kilian, A. (2006). DESIGN EXPLORATION WITH CIRCULAR DEPENDENCIES: A chair design experiment. ResearchGate. Retrieved February 5, 2020, from https://www.researchgate.net/publication/30871585_DESIGN_EXPLORATION_WITH_CIRCULAR_DEPENDENCIES_A_chair_design_experiment

124. Kim, T. H., & Cova, T. J. (2007). Tweening grammars: Deformation rules for representing change between discrete geographic entities. Computers, Environment and Urban Systems, GeoComputation 2005, 31(3), 317–336.

125. Kirkpatrick, D. G. (1979). Efficient computation of continuous skeletons. 20th Annual Symposium on Foundations of Computer Science (sfcs 1979) (pp. 18–27). Presented at the 20th Annual Symposium on Foundations of Computer Science (sfcs 1979).

126. Kitzbichler, M. G., Smith, M. L., Christensen, S. R., & Bullmore, E. (2009). Broadband Criticality of Human Brain Network Synchronization. PLOS Computational Biology, 5(3), e1000314.

127. Knight, T., & Stiny, G. (2015). Making grammars: From computing with shapes to computing with things. Design Studies, Special Issue: Computational Making, 41, 8–28.

128. Knight, T. W. (1980). The Generation of Hepplewhite-Style Chair-Back Designs. Environment and Planning B: Planning and Design, 7(2), 227–238.

129. Knight, T. W. (1981). The Forty-One Steps. Environment and Planning B: Planning and Design, 8(1), 97–114.

130. Knox, P. L. (2010). Cities and Design (1 edition.). London ; New York, NY: Routledge.

131. Koller, D., & Sahami, M. (1996). Toward Optimal Feature Selection. Proceedings of the Thirteenth International Conference on International Conference on Machine Learning, ICML'96 (pp. 284–292). San Francisco, CA,

USA: Morgan Kaufmann Publishers Inc. Retrieved August 7, 2019, from http://dl.acm.org/citation.cfm?id=3091696.3091731

132. Kostof, S. (1993). The City Shaped: Urban Patterns and Meanings Through History (Reprint edition.). New York; Boston; London: Bulfinch.

133. Kovács, I., Fehér, Á., & Julesz, B. (1998). Medial-point description of shape: A representation for action coding and its psychophysical correlates. Vision Research, 38(15), 2323–2333.

134. Krupic, J., Bauza, M., Burton, S., & O'Keefe, J. (2018). Local transformations of the hippocampal cognitive map. Science, 359(6380), 1143–1146.

135. Kruszelnicki, K. (2012). Curious and Curiouser. Retrieved August 19, 2019, from https://www.booktopia.com.au/curious-and-curiouser-dr-karl-kruszelnicki/prod9781742611709.html

136. KTU Politikos ir viešojo administravimo institutas. (2017). ESS8. Retrieved June 1, 2019, from http://www.lidata.eu/index_search_results_data.php?pid=LiDA%3Aquant.LiDA_ESS_0309

137. Lai-Kuan Wong, & Kok-Lim Low. (2009). Saliency-enhanced image aesthetics class prediction. 2009 16th IEEE International Conference on Image Processing (ICIP) (pp. 997–1000). Presented at the 2009 16th IEEE International Conference on Image Processing (ICIP).

138. Lang, J. (1987). Creating Architectural Theory: The Role of the Behavioral Sciences in Environmental Design. New York: Van Nostrand Reinhold.

139. Lawton, M. P., & Nahemow, L. (1973). Ecology and the aging process. The psychology of adult development and aging (pp. 619–674). Washington, DC, US: American Psychological Association.

140. Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324.

141. Lee, S. Y., & Tiong, K. M. (2013). Algorithmic Generation of Chinese Lattice Designs.

142. Leeds, D. D., Seibert, D. A., Pyles, J. A., & Tarr, M. J. (2013). Comparing visual representations across human fMRI and computational vision. Journal of Vision, 13(13), 25.

143. Lemarchand, G., & Lomberg, J. (1996). Are there any universal principles in science and aesthetics that could help us to set the unknown parameters for interstellar communication? ResearchGate. Retrieved February 5, 2020, from https://www.researchgate.net/publication/316427460_Are_there_any_universal_principles_in_science_and_aesthetics_that_could_help_us_to_set_the_unknown_parameters_for_interstellar_communication

144. Leyton, M. (1988). A Process-grammar for Shape. Artif. Intell., 34(2), 213–247.

145. Leyton, Michael. (1984). Perceptual organization as nested control. Biological Cybernetics, 51(3), 141–153.

146. Leyton, Michael. (1986a). A theory of information structure I. General principles. Journal of Mathematical Psychology, 30(2), 103–160.

147. Leyton, Michael. (1986b). Principles of Information Structure Common to Six Levels of the Human Cognitive System. Inf. Sci., 38(1), 1–120.

148. Leyton, Michael. (2006). Shape as Memory (1 edition.). Basel, Switzerland ; Boston: Birkhäuser Architecture.

149. Li, Y., & Hu, C.-J. (2010). Aesthetic Learning in an Interactive Evolutionary Art System. In C. Di Chio, A. Brabazon, G. A. Di Caro, M. Ebner, M. Farooq, A. Fink, J. Grahl, et al. (Eds.), Applications of Evolutionary Computation, Lecture Notes in Computer Science (pp. 301–310). Berlin, Heidelberg: Springer.

150. Licamele, L., & Getoor, L. (2006). Social Capital in Friendship-Event Networks. Sixth International Conference on Data Mining (ICDM’06) (pp. 959–964). Presented at the Sixth International Conference on Data Mining (ICDM’06).

151. Lindstrom, L., Dudfield, R., Shinners, P., Dudfield, N., & Kluyver, T. (2000). Pygame. Retrieved June 23, 2018, from https://www.pygame.org/wiki/about

152. Loffler, G. (2008). Perception of contours and shapes: Low and intermediate stage mechanisms. Vision Research, Vision Research Reviews, 48(20), 2106–2127.

153. Lorek, H., & White, M. (1993). Parallel Bird Flocking Simulation.

154. Losa, G. A. (2014). On the Fractal Design in Human Brain and Nervous Tissue. Applied Mathematics, 2014. Retrieved August 14, 2019, from http://www.scirp.org/journal/PaperInformation.aspx?PaperID=47368

155. Lowet, A. S., Firestone, C., & Scholl, B. J. (2018). Seeing structure: Shape skeletons modulate perceived similarity. Attention, Perception, & Psychophysics, 80(5), 1278–1289.

156. Lynch, K. (1960). The Image of the City. Cambridge, Mass.: The MIT Press.

157. Lynch, K. (1984). Good City Form (Reprint edition.). Cambridge, Mass: The MIT Press.

158. Macdonald, J., Stefanovic, D., & Stojanovic, M. (2009). Informatique. Des assemblages d’ADN rompus au jeu et au travail / J. Macdonald. Informatique. Des assemblages d’ADN rompus au jeu et au travail / J. Macdonald, Pour la science. Retrieved August 5, 2019, from https://mediatheque.ville-bourges.fr/NUMERIQUE/doc/SYRACUSE/1604278/informatique-des-assemblages-d-adn-rompus-au-jeu-et-au-travail-j-macdonald

159. MacLean, P. (1990). The Triune Brain in Evolution: Role in Paleocerebral Functions. Retrieved August 12, 2019, from https://www.bookdepository.com/Triune-Brain-Evolution-Paul-D-MacLean/9780306431685

160. Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related

structural change in the hippocampi of taxi drivers. Proceedings of the National Academy of Sciences of the United States of America, 97(8), 4398–4403.

161. Maguire, E. A., Nannery, R., & Spiers, H. J. (2006). Navigation around London by a taxi driver with bilateral hippocampal lesions. Brain: A Journal of Neurology, 129(Pt 11), 2894–2907.

162. Mandelbrot, B. (1967). How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. Science, 156(3775), 636–638.

163. Mandelbrot, B. B. (1982). The Fractal Geometry of Nature (2nd prt. edition.). San Francisco: Times Books.

164. Markowsky, G. (1992). Misconceptions about the Golden Ratio.

165. Marr, D., & Nishihara, K. (1987). Visual Information Processing: Artificial Intelligence and the Sensorium of Sight—ScienceDirect. Retrieved July 28, 2019, from https://www.sciencedirect.com/science/article/pii/B978008051581650060X

166. McIntosh, C. (Ed.). (2013). Cambridge Advanced Learner's Dictionary with CD-ROM (4 edition.). Cambridge: Cambridge University Press.

167. McKay, G. (2019). The Architecture of Architectures. Misfits' architecture. Retrieved August 26, 2020, from https://misfitsarchitecture.com/2019/07/14/the-architecture-of-architectures-2007/

168. Mede, T., Chambon, G., Hagenmuller, P., & Nicot, F. (2018). A medial axis based method for irregular grain shape representation in DEM simulations. ResearchGate. Retrieved July 25, 2019, from https://www.researchgate.net/publication/322799877_A_medial_axis_based_method_for_irregular_grain_shape_representation_in_DEM_simulations

169. Michael Leyton. (2018). . Retrieved December 6, 2018, from https://psych.rutgers.edu/faculty-profiles-a-contacts/122-michael-leyton

170. Milios, E. E. (1989). Shape Matching Using Curvature Processes. Comput. Vision Graph. Image Process., 47(2), 203–226.

171. Montello, D. (2007). The contribution of space syntax to a comprehensive theory of environmental psychology. Proceedings, 6.

172. Moughtin, J. C., MRTPI, R. C. Bs. M., MA, C. S. Bs., PhD, P. S. L. in U. U. of R. C. I., & BSc, R. C. M. (1999). Urban Design: Method and Technique. Oxford, England ; Boston: Architectural Press.

173. Moura, F. A., Martins, L. E. B., Anido, R. O., Ruffino, P. R. C., Barros, R. M. L., & Cunha, S. A. (2013). A spectral analysis of team dynamics and tactics in Brazilian football. Journal of Sports Sciences, 31(14), 1568–1577.

174. Murray, J., Blackburn, G. A., Whyatt, J. D., & Edwards, C. (2018). Using fractal analysis of crown images to measure the structural condition of trees. Forestry: An International Journal of Forest Research, 91(4), 480–491.

175. Musto, C., Semeraro, G., Lops, P., & Gemmis, M. de. (2015). CrowdPulse: A framework for real-time semantic analysis of social streams. Information Systems, 54, 127–146.

176. Naila, M., Marchesotti, L., & Perronnin, F. (2017). AVA: A Large-Scale Database for Aesthetic Visual Analysis. Academic Torrents. Retrieved

February 11, 2020, from http://academictorrents.com/details/71631f83b11d3d79d8f84efe0a7e12f0ac001460

177. Nakagaki, T., Yamada, H., & Tóth, Á. (2000). Maze-solving by an amoeboid organism. Nature, 407(6803), 470–470.

178. Neufert, E., & Neufert, P. (1970). Neufert Architects' Data. Retrieved January 29, 2020, from https://www.amazon.com/Neufert-Architects-Data-Fourth-Ernst/dp/1405192534

179. Newman, M. E. J. (2011). Complex Systems: A Survey. ResearchGate. Retrieved August 19, 2019, from https://www.researchgate.net/publication/277069336_Complex_Systems_A_Survey

180. Newman, M. E. J. (2013). Power laws, Pareto distributions and Zipf's law. Cities, 30, 59–67.

181. Ng, E. (Ed.). (2009). Designing High-Density Cities: For Social and Environmental Sustainability (1 edition.). London ; Sterling, VA: Routledge.

182. Novák, V., Perfilieva, I., & Mockor, J. (1999). Mathematical Principles of Fuzzy Logic. The Springer International Series in Engineering and Computer Science. Springer US. Retrieved June 24, 2018, from //www.springer.com/gp/book/9780792385950

183. NumPy. (2018). NumPy is the fundamental package for scientific computing with Python. Retrieved July 12, 2019, from https://www.numpy.org/

184. Obando, R. A. (2015). Partitioning of Cellular Automata Rule Spaces. Complex Systems, Volume 24, Issue 1. Retrieved August 16, 2019, from https://www.complex-systems.com/abstracts/v24_i01_a02/

185. Ogunsanya, A., & Ade, B. (1986). Graph theory in intra-urban traffic flow estimation. GeoJournal, 12(3), 334–336.

186. OpenStreetMap. (2019). OpenStreetMap. Retrieved May 31, 2019, from https://www.openstreetmap.org/

187. OSGeo. (2016). OSGeo. Retrieved August 6, 2019, from https://www.osgeo.org/

188. Osmosis. (2016). . Java, OpenStreetMap on GitHub. Retrieved August 6, 2019, from https://github.com/openstreetmap/osmosis

189. O'Sullivan, D. (2002). Toward micro-scale spatial modeling of gentrification. Journal of Geographical Systems, 4(3), 251–274.

190. Parish, Y. I. H., & Müller, P. (2001). Procedural modeling of cities. Proceedings of the 28th annual conference on Computer graphics and interactive techniques, SIGGRAPH '01 (pp. 301–308). New York, NY, USA: Association for Computing Machinery. Retrieved May 4, 2020, from https://doi.org/10.1145/383259.383292

191. Park, T., & Casella, G. (2008). The Bayesian Lasso. Journal of the American Statistical Association, 103(482), 681–686.

192. Pearson, J. E. (1993). Complex Patterns in a Simple System. Science, 261(5118), 189–192.

193. Petrović, N. (2010). Medial axis extraction and thickness measurement of formed sheet metal parts.

194. Pingala. (200AD). Chandasshastra.

195. Pinto, N. N., & Antunes, A. P. (2010). A Cellular Automata Model Based on Irregular Cells: Application to Small Urban Areas. Environment and Planning B: Planning and Design, 37(6), 1095–1114.

196. Poli, R., Langdon, W. B., & McPhee, N. F. (2008). A Field Guide to Genetic Programming. Lulu Enterprises, UK Ltd.

197. PostGIS, S. and G. O. for P. (2018a). PostGIS — Spatial and Geographic Objects for PostgreSQL. Retrieved July 12, 2019, from https://postgis.net/

198. PostGIS, S. and G. O. for P. (2018b). PostGIS — Spatial and Geographic Objects for PostgreSQL. Retrieved July 12, 2019, from https://postgis.net/

199. Prisms. (2012, February 27). Fxguide. Retrieved December 18, 2018, from https://www.fxguide.com/featured/side-effects-software-25-years-on/

200. Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. Genetics, 155(2), 945–959.

201. Prusinkiewicz, P., & Lindenmayer, A. (1990). The Algorithmic Beauty of Plants. The Virtual Laboratory. New York: Springer-Verlag. Retrieved August 8, 2019, from https://www.springer.com/gp/book/9780387946764

202. Psycopg. (2018). Psycopg is the most popular PostgreSQL adapter for the Python programming language. Retrieved July 12, 2019, from http://initd.org/psycopg/

203. Pullum, G. K., & Kornai, A. (2003). Mathematical linguistics.

204. Putnam, R. D. (2001a). Bowling Alone: The Collapse and Revival of American Community. Retrieved June 1, 2019, from http://bowlingalone.com/

205. Putnam, R. D. (2001b). Bowling Alone: The Collapse and Revival of American Community. Retrieved June 1, 2019, from http://bowlingalone.com/

206. QGIS. (2017a). . Retrieved May 10, 2017, from http://www.qgis.org/en/site/

207. QGIS. (2017b). . Retrieved May 10, 2017, from http://www.qgis.org/en/site/

208. Reid, C. R., Latty, T., Dussutour, A., & Beekman, M. (2012). Slime mold uses an externalized spatial "memory" to navigate in complex environments. Proceedings of the National Academy of Sciences, 109(43), 17490–17494.

209. Rekvizitai.lt. (2019). . Retrieved May 31, 2019, from https://rekvizitai.vz.lt/

210. Riddell, J. B., & Harvey, M. E. (1972). The urban system in the migration process: An evaluation of step-wise migration in Sierra Leone. Economic geography.

211. Romeijn, J.-W. (2017). Philosophy of Statistics. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Spring 2017.). Metaphysics Research Lab, Stanford University.

212. Rosenfeld, A., & Pfaltz, J. L. (1966). Sequential Operations in Digital Picture Processing. J. ACM, 13, 471–494.

213. Rossum, G. V., & Drake, F. L. J. (2011a). The Python Language Reference Manual. Bristol: Network Theory Ltd.

214. Rossum, G. V., & Drake, F. L. J. (2011b). The Python Language Reference Manual. Bristol: Network Theory Ltd.

215. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). Learning Internal Representations by Error Propagation (No. ICS-8506). CALIFORNIA UNIV SAN DIEGO LA JOLLA INST FOR COGNITIVE SCIENCE. Retrieved August 7, 2019, from https://apps.dtic.mil/docs/citations/ADA164453

216. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533–536.

217. Rupasingha, A., Goetz, S. J., & Freshwater, D. (2006). The production of social capital in US counties. The Journal of Socio-Economics, 35(1), 83–101.

218. Saarinen, E., Pelkonen, E.-L., & Albrecht, D. (2006). Eero Saarinen: Shaping the future. New Haven : Yale University Press ; New York : In association with The Finnish Cultural Institute in New York ; Helsinki : The Museum of Finnish Architecture ; Washington, D.C. : The National Building Museum ; New Haven : Yale University School of Architecture. Retrieved August 14, 2019, from https://trove.nla.gov.au/version/46579340

219. Sakuragawa, S., Miyazaki, Y., Kaneko, T., & Makita, T. (2005). Influence of wood wall panels on physiological and psychological responses. Journal of Wood Science, 51(2), 136–140.

220. Salingaros, N. (2003). Connecting the Fractal City. Planum association. Retrieved May 5, 2020, from http://www.planum.net/connecting-the-fractal-city

221. Salingaros, N. (2020). The Biophilic Healing Index Predicts Effects of the Built Environment on Our Wellbeing. Retrieved August 26, 2020, from /paper/The-Biophilic-Healing-Index-Predicts-Effects-of-the-Salingaros/1ad2730cbc808207fa24506b3b69c23233fb39d2

222. Salingaros, N. A., & West, B. J. (1999). A Universal Rule for the Distribution of Sizes. Environment and Planning B: Planning and Design, 26(6), 909–923.

223. Sampson, R. J., & Graif, C. (2009). Neighborhood Social Capital as Differential Social Organization: Resident and Leadership Dimensions. American Behavioral Scientist, 52(11), 1579–1605.

224. Samuel, A. (1959). Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development, 3(3), 210–229. Presented at the IBM Journal of Research and Development.

225. Sanderson, K. (2008). Millennium Bridge wobble explained. Nature. Retrieved August 19, 2019, from http://www.nature.com/articles/news.2008.1311

226. Sapin, E. (2010). Gliders and Glider Guns Discovery in Cellular Automata. In A. Adamatzky (Ed.), Game of Life Cellular Automata (pp. 135–165).

London: Springer London. Retrieved August 16, 2019, from https://doi.org/10.1007/978-1-84996-217-9_9

227. Schaffranek, R. (2013). SPACE SYNTAX FOR GENERATIVE DESIGN : On the application of a new tool. Retrieved August 20, 2020, from /paper/SPACE-SYNTAX-FOR-GENERATIVE-DESIGN-%3A-On-the-of-a-Schaffranek/2c217f1229be7acd371e6ab645ed341771524ec5

228. Schelling, T. (1969a). Models of Segregation. American Economic Review, 59(2), 488–93.

229. Schelling, T. (1969b). Models of Segregation. American Economic Review, 59(2), 488–93.

230. Schroeder, M. (1991). Fractals, chaos, power laws. AIChE Journal, 38(1), 159–160.

231. Schwalb, M., & Roth, R. (1970). Axenic Growth and Development of the Cellular Slime Mould, Dictyostelium discoideum. ResearchGate. Retrieved August 7, 2019, from https://www.researchgate.net/publication/240256849_Axenic_Growth_and_Development_of_the_Cellular_Slime_Mould_Dictyostelium_discoideum

232. Semboloni, F. (1996). An Urban and Regional Model Based on Cellular Automata: Environment and Planning B: Planning and Design. SAGE PublicationsSage UK: London, England. Retrieved July 10, 2020, from https://journals.sagepub.com/doi/10.1068/b240589

233. SFCGAL. (2016). . Retrieved August 6, 2019, from http://www.sfcgal.org/

234. Shaftoe, H. (2012). Convivial Urban Spaces: Creating Effective Public Places. Earthscan.

235. Shamir, L. (2015). What makes a Pollock Pollock: A machine vision approach. International Journal of Arts and Technology, 8(1), 1.

236. Shvets, A. (2007). Design Patterns and Refactoring. Retrieved June 23, 2018, from https://sourcemaking.com/

237. Siegert, F. (1991). Analysis of optical density wave propagation and cell movement in the cellular slime mould Dictyostelium discoideum. Retrieved August 7, 2019, from https://www.researchgate.net/publication/222488878_Analysis_of_optical_density_wave_propagation_and_cell_movement_in_the_cellular_slime_mould_Dictyostelium_discoideum

238. Smith, M. S., Purser, N., & Giraud-Carrier, C. G. (2007). Social Capital in the Blogosphere: A Case Study. AAAI Spring Symposium: Social Information Processing.

239. Soja, E. (2000). Postmetropolis: Critical Studies of Cities and Regions. Wiley.com. Retrieved May 5, 2020, from https://www.wiley.com/en-us/Postmetropolis%3A+Critical+Studies+of+Cities+and+Regions-p-9781577180012

240. Sonka, M., Hlavac, V., & Boyle, R. (1993). Image Processing, Analysis and Machine Vision. Springer US. Retrieved August 14, 2019, from https://www.springer.com/gp/book/9780412455704

241. stat.gov.lt. (2011). Lietuvos Respublikos 2011 metų visuotinio gyventojų ir būstų surašymo eigos aprašas, 7.

242. Sternberg, R. J. (2003). Cognitive Psychology,. Retrieved August 19, 2019, from https://www.abebooks.com/Cognitive-Psychology-Third-Edition-Robert-Sternberg/3261609185/bd

243. Stiny, G. (1977). Ice-Ray: A Note on the Generation of Chinese Lattice Designs. Retrieved February 10, 2020, from https://journals.sagepub.com/doi/10.1068/b040089

244. Stiny, G. (1980). Kindergarten Grammars: Designing with Froebel's Building Gifts.

245. Stiny, G., & Gips, J. (1971). Shape Grammars and the Generative Specification of Painting and Sculpture.

246. Stiny, G., & Mitchell, W. (1980). The Grammar of Paradise: On the Generation of Mughul Gardens. Retrieved February 10, 2020, from https://journals.sagepub.com/doi/abs/10.1068/b070209

247. Stiny, G., & Mitchell, W. J. (1978). The Palladian Grammar. Environment and Planning B: Planning and Design, 5(1), 5–18.

248. Subbian, K., Sharma, D., Wen, Z., & Srivastava, J. (2013). Finding influencers in networks using social capital. 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013) (pp. 592–599). Presented at the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013).

249. Taylor, R. P., & Spehar, B. (2016). Fractal Fluency: An Intimate Relationship Between the Brain and Processing of Fractal Stimuli. Retrieved August 14, 2019, from https://www.growkudos.com/publications/10.1007%252F978-1-4939-3995-4_30

250. Taylor, R. P., Spehar, B., Van Donkelaar, P., & Hagerhall, C. M. (2011). Perceptual and Physiological Responses to Jackson Pollock's Fractals. Frontiers in Human Neuroscience, 5. Retrieved August 14, 2019, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3124832/

251. Thompson, D. (1942). On Growth and Form: The Complete Revised Edition (Revised edition.). New York: Dover Publications.

252. Tiwari, S., Sharma, V., Mujawar, M., Mishra, Y. K., Kaushik, A., & Ghosal, A. (2019). Biosensors for Epilepsy Management: State-of-Art and Future Aspects. Sensors (Basel, Switzerland), 19(7). Retrieved August 19, 2019, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6480455/

253. Tobler, W. (1975). Cellular Geography. Monograph, IIASA, Laxenburg, Austria: WP-75-100. Retrieved July 7, 2020, from http://pure.iiasa.ac.at/id/eprint/328/

254. Touboul, J., & Destexhe, A. (2010). Can Power-Law Scaling and Neuronal Avalanches Arise from Stochastic Dynamics? PLOS ONE, 5(2), e8982.

255. Tsuda, S., Artmann, S., & Zauner, K.-P. (2009). The Phi-Bot: A Robot Controlled by a Slime Mould. In A. Adamatzky & M. Komosinski (Eds.), Artificial Life Models in Hardware (pp. 213–232). Springer-Verlag London. Retrieved August 7, 2019, from https://eprints.soton.ac.uk/268247/

256. Turing, A. M. (1952). The chemical basis of morphogenesis. Phil. Trans. R. Soc. Lond. B, 237(641), 37–72.

257. Ulrich, R. S. (1984). View through a window may influence recovery from surgery. Science (New York, N.Y.), 224(4647), 420–421.

258. US Census Bureau. (2016). Cartographic Boundary Shapefiles—Counties. Retrieved May 9, 2017, from https://www.census.gov/geo/maps-data/data/cbf/cbf_counties.html

259. Vasileiou, V., Granot, J., Piran, T., & Amelino-Camelia, G. (2015). A Planck-scale limit on spacetime fuzziness and stochastic Lorentz invariance violation. Nature Physics, 11, 344–346.

260. Vitruvius, P. (15ADa). De architectura. Torino: Giulio Einaudi.

261. Vitruvius, P. (15ADb). De architectura. Torino: Giulio Einaudi.

262. Von Neumann, J., & Burks, A. W. (Arthur W. (1966). Theory of self-reproducing automata. Urbana, University of Illinois Press. Retrieved June 23, 2018, from http://archive.org/details/theoryofselfrepr00vonn_0

263. Voronoi, G. (1908). Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Premier mémoire. Sur quelques propriétés des formes quadratiques positives parfaites. Journal für die reine und angewandte Mathematik, 133, 97–178.

264. vrk.lt. (2016). Atviri duomenys—Vrk.lt. Retrieved May 31, 2019, from https://www.vrk.lt/atviri-duomenys

265. Wade, A. (2010). Game of Life News: Oblique Life spaceship created. Retrieved August 15, 2019, from http://pentadecathlon.com/lifeNews/2010/05/oblique_life_spaceship_created.html

266. Wang, K., & Taylor, R. B. (2006). Simulated walks through dangerous alleys: Impacts of features and progress on fear. Journal of environmental psychology, 26(4), 269–283.

267. Wang, X., & Burbeck, C. A. (1998). Scaled medial axis representation: Evidence from position discrimination task. Vision Research, 38(13), 1947–1959.

268. Webster, C. J., & Wu, F. (1999). Regulation, Land-Use Mix, and Urban Performance. Part 1: Theory: Environment and Planning A. SAGE PublicationsSage UK: London, England. Retrieved July 18, 2020, from https://journals.sagepub.com/doi/10.1068/a311433

269. Weier, H. U., & Gray, J. W. (1988). A programmable system to perform the polymerase chain reaction. DNA (Mary Ann Liebert, Inc.), 7(6), 441–447.

270. White, R., & Engelen, G. (1993). Cellular Automata and Fractal Urban Form: A Cellular Modelling Approach to the Evolution of Urban Land-Use Patterns: Environment and Planning A. SAGE PublicationsSage UK: London,

England. Retrieved July 10, 2020, from https://journals.sagepub.com/doi/10.1068/a251175

271. Whyte, W. H. (2001). The Social Life of Small Urban Spaces (unknown edition.). New York, NY: Project for Public Spaces.

272. Wilder, J., Feldman, J., & Singh, M. (2011a). Superordinate shape classification using natural shape statistics. Cognition, 119(3), 325–340.

273. Wilder, J., Feldman, J., & Singh, M. (2011b). Superordinate shape classification using natural shape statistics. Cognition, 119(3), 325–340.

274. Wilder, J., Feldman, J., & Singh, M. (2011c). Superordinate shape classification using natural shape statistics. Cognition, 119(3), 325–340.

275. Wilson, E. O. (1984). Biophilia. Retrieved August 5, 2019, from https://portals.iucn.org/library/node/27612

276. Wolfram, S. (1983). Statistical mechanics of cellular automata. Reviews of Modern Physics, 55(3), 601–644.

277. Wolfram, S. (2002). A New Kind of Science (1 edition.). Champaign, IL: Wolfram Media.

278. Woolf, G. (2012). Divinity and Power in ancient Rome. RELIGION AND POWER. Divine Kingship in the Ancient world and beyond. Retrieved August 14, 2019, from https://www.academia.edu/554147/Divinity_and_Power_in_ancient_Rome

279. Wu, X., & Foster, D. J. (2014). Hippocampal Replay Captures the Unique Topological Structure of a Novel Environment. Journal of Neuroscience, 34(19), 6459–6469.

280. Wu, Y., Bauckhage, C., & Thurau, C. (2010). The Good, the Bad, and the Ugly: Predicting Aesthetic Image Labels. 2010 20th International Conference on Pattern Recognition (pp. 1586–1589). Presented at the 2010 20th International Conference on Pattern Recognition.

281. Yang, P. P.-J., Putra, S. Y., & Chaerani, M. (2007). Computing the sense of time in urban physical environment. URBAN DESIGN International, 12(2), 115–129.

282. Yates, F. (2014). The Art of Memory. London: Random House UK.

283. Yosinski, J., Clune, J., Nguyen, A., Fuchs, T., & Lipson, H. (2015). Understanding Neural Networks Through Deep Visualization. ArXiv:1506.06579 [cs]. Retrieved July 22, 2019, from http://arxiv.org/abs/1506.06579

284. Zaleckis, K. (2018). Erdvės sintaksė urbanistinei analizei: Koncepcijos, apskaičiavimai ir pavyzdžiai.

285. Менделеев, Д. И. (1869). Соотношение свойств с атомным весом элементов. OZON.ru, Журнал Русского Химического общества. Том II за 1870 год — купить в интернет-магазине OZON с быстрой доставкой. Retrieved August 16, 2019, from https://www.ozon.ru/context/detail/id/32795264/

6. CURRICULUM VITAE

Marius Ivaškevičius
Marius.Ivaskevicius@ktu.lt

Education:
1999–2001 Kaunas College, Furniture Restoration
2001–2004 Art Academy of Vilnius, Faculty of Kaunas, Bachelor's degree in Architecture
2004–2006 Vilnius Gediminas Technical University, Master's degree in Architecture
2015–2021 Kaunas University of Technology, Doctor of Philosophy (PhD) in Arts, Humanities

Professional experience:
2003–2004 Architect at LLC "Romos Projektas"
2004–2005 Quality Assurance at LLC "Orbis Avia"
2007–2009 Urbanist at LLC "GB Technologijos"
2009–2015 Architect at Kaunas MC "Kauno planas"
2016–2017 Architect at LCC "JAS"
2019–2020 Data Analyst at LCC "Edukacinės sistemos"
2019–now Lecturer at KTU SAF

Research interest:
Interests: urbanism, social capital, parametric architecture, parametric urbanism.
Methods: statistical analysis, machine learning, deep neural networks, agent modeling, cellular automata.

Publications related to the thesis:
1. Using Natural Shape Statistics of Urban Form to Model Social Capital. In *Science – Future of Lithuania*, Vol 9, No. 1, (2017).
2. Lietuvos socialinio kapitalo modeliavimas naudojant formos skeleto metodą, In *Science – Future of Lithuania*, Vol 12, (2020).
3. Biophilic Analysis of Façades in Kaunas City Centre in Landscape Architecture and Art at Latvia University of Life Sciences and Technologies.

Scientific conferences:
1. Using Natural Shape Statistics of Urban Form to Model Social Capital at K. Šešelgis' readings – 2017.
2. Urban Form (as Memory) Influence to Social Capital at Second Baltic Conference for Young Researchers at Architecture and Urban Planning at Riga Technical University.

3. Urban Form (as Memory) Influence to Social Capital at the 2nd Baltic Conference of Young Researchers in Architecture, Landscape Architecture and Urbanism at Vilnius Tech.

7. ANNEXES

Annex 1. Social capital index of Lithuania

name	votep	social	nonProfit	sci
Akmenės rajono savivaldybė	0.500461851	10.72983727	6	3.272121
Tauragės rajono savivaldybė	0.476759592	6.584800124	7	-1.01956
Lazdijų rajono savivaldybė	0.523117162	5.925783558	4	-1.5839
Molėtų rajono savivaldybė	0.5	7.452468839	5	0.073486
Alytaus miesto savivaldybė	0.486245301	9.010520923	15	1.592938
Kalvarijos savivaldybė	0.436482939	5.61069514	2	-1.93825
Šiaulių rajono savivaldybė	0.484611881	5.240713689	7	-2.31103
Mažeikių rajono savivaldybė	0.459501282	5.952273582	11	-1.54088
Rokiškio rajono savivaldybė	0.51284719	7.213886551	8	-0.18796
Šalčininkų rajono savivaldybė	0.543758765	4.008901115	4	-3.62087
Skuodo rajono savivaldybė	0.461722324	7.087853334	4	-0.40232
Visagino savivaldybė	0.359274123	13.88271573	4	6.046788
Alytaus rajono savivaldybė	0.502149122	4.587644485	1	-3.31449
Elektrėnų savivaldybė	0.430691136	7.449097824	8	0.275724
Kupiškio rajono savivaldybė	0.502167814	8.268789179	3	0.567222
Jonavos rajono savivaldybė	0.458932349	6.938904709	8	-0.64305
Plungės rajono savivaldybė	0.475966035	7.550821645	4	-0.26391
Varėnos rajono savivaldybė	0.527589545	8.783860356	3	0.964516
Kauno miesto savivaldybė	0.535451779	11.46665102	134	4.450568
Šilalės rajono savivaldybė	0.514408308	5.897569927	8	-1.23606
Vilkaviškio rajono savivaldybė	0.474100803	5.067781573	7	-2.44831
Kretingos rajono savivaldybė	0.456912197	6.399830076	8	-1.10438
Šakių rajono savivaldybė	0.485062357	7.658791662	11	0.56854
Klaipėdos rajono savivaldybė	0.490707427	5.804672755	7	-1.8728
Kazlų Rūdos savivaldybė	0.455844891	7.91491026	2	0.22669
Švenčionių rajono savivaldybė	0.528185988	6.869111026	2	-1.01222
Kauno rajono savivaldybė	0.542152704	5.580946341	21	-1.78249
Joniškio rajono savivaldybė	0.464105549	8.260875644	10	1.310486
Ignalinos rajono savivaldybė	0.47208695	7.999788427	4	0.533637
Druskininkų savivaldybė	0.524625935	9.916806028	6	2.517775
Telšių rajono savivaldybė	0.469811123	6.779695311	10	-0.66412
Ukmergės rajono savivaldybė	0.511634357	7.704554407	7	0.111128

Kėdainių rajono savivaldybė	0.529734614	7.208794722	5	-0.62235
Klaipėdos miesto savivaldybė	0.440007335	10.43216635	59	3.266521
Pagėgių savivaldybė	0.460645526	5.357211183	2	-2.02016
Anykščių rajono savivaldybė	0.510294475	8.46183962	6	0.94519
Utenos rajono savivaldybė	0.536455279	8.487005808	6	0.740667
Radviliškio rajono savivaldybė	0.487548835	6.346279167	8	-1.14481
Šiaulių miesto savivaldybė	0.525254975	10.21122681	34	2.902883
Palangos miesto savivaldybė	0.51334789	14.95545872	11	8.573209
Panevėžio rajono savivaldybė	0.48519011	3.975406279	2	-3.85156
Kaišiadorių rajono savivaldybė	0.482680722	6.092572277	9	-1.16447
Širvintų rajono savivaldybė	0.517975567	8.639632092	3	0.989025
Trakų rajono savivaldybė	0.453213313	7.576437272	12	0.478248
Pasvalio rajono savivaldybė	0.501709851	8.011292099	2	0.07665
Pakruojo rajono savivaldybė	0.498793771	8.196087519	2	0.301246
Raseinių rajono savivaldybė	0.585689165	7.55275024	10	0.244157
Jurbarko rajono savivaldybė	0.485479693	8.206487851	7	0.762101
Neringos savivaldybė	0.359274123	21.61244257	1	13.87713
Vilniaus miesto savivaldybė	0.52477392	16.73548835	325	9.940297
Vilniaus rajono savivaldybė	0.53771841	3.799762946	7	-3.96322
Kelmės rajono savivaldybė	0.473855421	6.728539148	8	-0.58867
Prienų rajono savivaldybė	0.485383342	6.426709873	9	-0.73646
Zarasų rajono savivaldybė	0.429431476	11.54485861	4	3.917353
Birštono savivaldybė	0.613290632	13.98103661	3	7.557347
Panevėžio miesto savivaldybė	0.527576079	8.729975318	26	1.334439
Rietavo savivaldybė	0.460199005	8.64624505	3	1.536049
Šilutės rajono savivaldybė	0.462542582	7.573221082	7	-0.0791
Biržų rajono savivaldybė	0.499434288	8.096541612	2	0.160586
Marijampolės savivaldybė	0.480928359	7.716714242	10	0.072602

Annex 2. Graph analysis code

createTemplateTable.sql

```
-- function anaskel(integer,geometry)
DROP FUNCTION IF EXISTS anaskel(int,geometry);
DROP TABLE IF EXISTS results0;
DROP TABLE IF EXISTS results;
CREATE TABLE results
(
  id integer,
-- len
  len integer,
        lenpruned integer,
--      depth
  depthmax integer,
        depthmaxpruned integer,
  depthmean integer,
  depthmeanpruned integer,
-- touchingn
  touchingnmax integer,
        touchingnmaxpruned integer,
  touchingnmean integer,
  touchingnmeanpruned integer,
-- distance
        distancemin numeric,
        distanceminpruned numeric,
        distancemax numeric,
        distancemaxpruned numeric,
  distancemean numeric,
  distancemeanpruned numeric,
        distancesum numeric,
        distancesumpruned numeric,
-- azimuth
        azimuthmean numeric,
  azimuthmeanpruned numeric,
-- turnangle
        turnanglemin numeric,
        turnangleminpruned numeric,
        turnanglemax numeric,
        turnanglemaxpruned numeric,
  turnanglemean numeric,
  turnanglemeanpruned numeric,
        turnanglesum numeric,
        turnanglesumpruned numeric,
-- turnangleabs
```

```
          turnangleabsmin numeric,
          turnangleabsminpruned numeric,
          turnangleabsmax numeric,
          turnangleabsmaxpruned numeric,
   turnangleabsmean numeric,
   turnangleabsmeanpruned numeric,
          turnangleabssum numeric,
          turnangleabssumpruned numeric,
-- wigglines
          wigglines numeric,
          wigglinespruned numeric,
-- nrofbranches
          nrofbranches integer,
          nrofbranchespruned integer,
-- branchdepth
          branchdepthmax integer,
          branchdepthmaxpruned integer,
   branchdepthmean integer,
   branchdepthmeanpruned integer
);
```

updatetouches.sql

```
CREATE OR REPLACE FUNCTION updatetouches()RETURNS VOID AS $$
BEGIN
          -- add field with number of touching lines
          ALTER TABLE skeletonmp DROP COLUMN IF EXISTS touchingn;
      ALTER TABLE skeletonmp ADD touchingn INT DEFAULT 0;
      UPDATE skeletonmp AS s
         SET touchingn = j.count
         FROM(
            SELECT a.id AS id ,count(b.id) AS count
               FROM skeletonmp AS a
               INNER JOIN skeletonmp AS b
               ON ST_Touches(a.geom, b.geom)
               WHERE a.id!=b.id
               GROUP BY a.id
            ) AS j
         WHERE s.id = j.id;
      -- add field with array of touching lines ids for every line
      ALTER TABLE skeletonmp DROP COLUMN IF EXISTS touchingarr;
      ALTER TABLE skeletonmp ADD touchingarr INT ARRAY;
      UPDATE skeletonmp AS s
         SET touchingarr = j.arr
         FROM(
```

```
    SELECT a.id AS id ,array_agg(b.id) AS arr
      FROM skeletonmp AS a
      INNER JOIN skeletonmp AS b
      ON ST_Touches(a.geom, b.geom)
      WHERE a.id!=b.id
      GROUP BY a.id
    ) AS j
  WHERE s.id = j.id;
-- add field with number of start point touching lines
ALTER TABLE skeletonmp DROP COLUMN IF EXISTS touchingstartn;
ALTER TABLE skeletonmp ADD touchingstartn INT DEFAULT 0;
UPDATE skeletonmp AS s
  SET touchingstartn = j.count
  FROM(
    SELECT a.id AS id ,count(b.id) AS count
      FROM skeletonmp AS a
      INNER JOIN skeletonmp AS b
      ON ST_Touches(ST_StartPoint(a.geom), b.geom)
      WHERE a.id!=b.id
      GROUP BY a.id
    ) AS j
  WHERE s.id = j.id;
-- add field with number of end point touching lines
ALTER TABLE skeletonmp DROP COLUMN IF EXISTS touchingendn;
ALTER TABLE skeletonmp ADD touchingendn INT DEFAULT 0;
UPDATE skeletonmp AS s
  SET touchingendn = j.count
  FROM(
    SELECT a.id AS id ,count(b.id) AS count
      FROM skeletonmp AS a
      INNER JOIN skeletonmp AS b
      ON ST_Touches(ST_EndPoint(a.geom), b.geom)
      WHERE a.id!=b.id
      GROUP BY a.id
    ) AS j
  WHERE s.id = j.id;
-- add field with array of start point touching lines
ALTER TABLE skeletonmp DROP COLUMN IF EXISTS touchingstartarr;
ALTER TABLE skeletonmp ADD touchingstartarr INT ARRAY;
UPDATE skeletonmp AS s
  SET touchingstartarr = j.count
  FROM(
    SELECT a.id AS id ,array_agg(b.id) AS count
      FROM skeletonmp AS a
```

```
                INNER JOIN skeletonmp AS b
                ON ST_Touches(ST_StartPoint(a.geom), b.geom)
                WHERE a.id!=b.id
                GROUP BY a.id
            ) AS j
        WHERE s.id = j.id;
    -- add field with array of end point touching lines
    ALTER TABLE skeletonmp DROP COLUMN IF EXISTS touchingendarr;
    ALTER TABLE skeletonmp ADD touchingendarr INT ARRAY;
    UPDATE skeletonmp AS s
        SET touchingendarr = j.count
        FROM(
            SELECT a.id AS id ,array_agg(b.id) AS count
                FROM skeletonmp AS a
                INNER JOIN skeletonmp AS b
                ON ST_Touches(ST_EndPoint(a.geom), b.geom)
                WHERE a.id!=b.id
                GROUP BY a.id
            ) AS j
        WHERE s.id = j.id;
END;
$$ LANGUAGE plpgsql;
```

findDepth.sql

```
CREATE OR REPLACE FUNCTION findDepth()
RETURNS integer AS $finalDepth$
declare
        finalDepth integer;
    conttest boolean;
    maxdepth integer;
BEGIN
        ALTER TABLE skeletonmp DROP COLUMN IF EXISTS depth; -- erases
data
    ALTER TABLE skeletonmp DROP COLUMN IF EXISTS depthold; -- erases
data
    SELECT count(*) INTO maxdepth FROM skeletonmp;
        ALTER TABLE skeletonmp ADD depth INT DEFAULT 0;
    ALTER TABLE skeletonmp ADD depthold INT DEFAULT 0;
    UPDATE skeletonmp SET depth = maxdepth;
        UPDATE skeletonmp SET depth = 0 WHERE touchingstartn = 0 OR
touchingendn = 0;
        SELECT bool_or(depth<>depthold) INTO conttest FROM skeletonmp;
    WHILE conttest LOOP
        UPDATE skeletonmp SET depthold=depth;
```

```
        UPDATE skeletonmp AS s
            SET depth = j.bdepth + 1
            FROM(
                SELECT a.id AS id ,min(b.depth) AS bdepth
                    FROM skeletonmp AS a
                    INNER JOIN skeletonmp AS b
                    -- ON ST_Touches(a.geom, b.geom)
                          ON a.touchingarr @> ARRAY[b.id]
                    WHERE a.id!=b.id
                    GROUP BY a.id
                ) AS j
            WHERE s.id = j.id AND depth > j.bdepth + 1;
                          SELECT bool_or(depth<>depthold) INTO conttest FROM
skeletonmp;
             END LOOP;
   SELECT max(depth) INTO finalDepth FROM skeletonmp;
   RETURN finalDepth;
END;
$finalDepth$ LANGUAGE plpgsql;
-- tests

-- SELECT findDepth();
```

popOff.sql

```
CREATE OR REPLACE FUNCTION pop_off(arr ANYARRAY) RETURNS
ANYARRAY AS $$
             DECLARE
                          arrb INT[] := NULL;
    BEGIN
             IF array_upper(arr,1) > 1 THEN
             RETURN (SELECT arr[2:array_upper(arr,1)]);
        ELSE
             -- RETURN ARRAY[]::INT;
            RETURN arrb;
        END IF;
    END;
$$ LANGUAGE plpgsql;
-- tests
-- SELECT pop_off(ARRAY[ARRAY[1,2],ARRAY[3,4]]);
-- SELECT pop_off(ARRAY[ARRAY[1,2]]);
```

popValue.sql

```
CREATE OR REPLACE FUNCTION popvalue(val INT,arr INT[]) RETURNS
INT[] AS $$
```

```
          DECLARE
                    arrn INT[] := NULL;
      testval INT;
   BEGIN
          IF arr IS NOT NULL THEN
         IF array_upper(arr,1) = 1 AND arr[1] = val THEN
            RETURN NULL;
         ELSE
            FOREACH testval IN ARRAY arr LOOP
               IF testval <> val THEN
                  IF arrn IS NULL THEN
                     arrn := ARRAY[testval];
                  ELSE
                     arrn := arrn || ARRAY[testval];
                  END IF;
               END IF;
            END LOOP;
            RETURN arrn;
         END IF;
      ELSE
                              RETURN arrn;
                    END IF;
   END;
$$ LANGUAGE plpgsql;
-- tests
-- SELECT pop_off(ARRAY[ARRAY[1,2],ARRAY[3,4]]);
-- SELECT popvalue(2,ARRAY[1,2,3]);
-- SELECT popvalue(2,ARRAY[2]);
```

getdirbrances.sql

```
-- DROP FUNCTION getdirbrances(integer,boolean);
CREATE OR REPLACE FUNCTION getdirbrances(getid INT,dir BOOLEAN)
RETURNS INT[] AS $$
   DECLARE
      nextcrossroad INT[];
          BEGIN
          IF dir THEN
                    SELECT touchingstartarr INTO nextcrossroad FROM skeletonmp
WHERE id = getid;
                    ELSE
          SELECT touchingendarr INTO nextcrossroad FROM skeletonmp WHERE
id = getid;
      END IF;
```

```
        RETURN nextcrossroad;
            END;
            $$ LANGUAGE plpgsql;
-- test
-- SELECT getdirbrances(56,0::BOOLEAN);
```

solvenode.sql

```
CREATE OR REPLACE FUNCTION solvenode(unsolvednodes INT[]) RETURNS
INT[] AS $$
   DECLARE
            unsolvednode INT[];
        unsolvednodestmp INT[];
        curnodeid INT;
        dir INT;
            curbranchazimuth NUMERIC;
        lastbranchdepth INT;
        lastbranchid INT;
        thisbranchid INT;
        maxbranchid INT;
            nextcrossroad INT[];
        nexttrunkangle NUMERIC;
        nexttrunkid INT;
        nextbranchid INT;
        otherbranchids INT[];
        tmpid INT;
        iteration INT;
            BEGIN
        curnodeid := unsolvednodes[1][1];
        dir := unsolvednodes[1][2];
        unsolvednodes := pop_off(unsolvednodes);
        -- fill angles of selected dirrection
        SELECT azimuth INTO curbranchazimuth FROM skeletonmp WHERE id =
curnodeid;
        SELECT branchdepth INTO lastbranchdepth FROM skeletonmp WHERE id =
curnodeid;
        SELECT branchid INTO lastbranchid FROM skeletonmp  WHERE id =
curnodeid;
        SELECT max(branchid) INTO maxbranchid FROM skeletonmp;
        -- SELECT max(walkeriteration) INTO iteration FROM skeletonmp;
        -- UPDATE skeletonmp SET walkeriteration = iteration + 1 WHERE id =
curnodeid;
        -- SELECT touchingstartarr INTO nextcrossroad FROM skeletonmp WHERE
id = maxdepthid;
        -- use getdirbrances(getid INT,dir BOOLEAN) instead
```

```
        SELECT getdirbrances(curnodeid,dir::BOOLEAN) INTO nextcrossroad;
        UPDATE skeletonmp SET turnangle = azimuth - curbranchazimuth WHERE
ARRAY[id] <@ nextcrossroad;
        -- fill turnangleabs of selected dirrection
        UPDATE skeletonmp SET turnangleabs = abs(azimuth - curbranchazimuth)
WHERE ARRAY[id] <@ nextcrossroad;
        -- get trunk branch id
        SELECT min(turnangleabs) INTO nexttrunkangle FROM skeletonmp WHERE
ARRAY[id] <@ nextcrossroad;
        SELECT id INTO nexttrunkid FROM skeletonmp WHERE turnangleabs =
nexttrunkangle LIMIT 1;
        -- fill branchid & branchdepth of selected dirrection of trunk branch
        SELECT branchid INTO thisbranchid FROM skeletonmp WHERE id =
curnodeid;
        UPDATE skeletonmp SET branchid = thisbranchid WHERE id = nexttrunkid;
        UPDATE skeletonmp SET branchdepth = lastbranchdepth WHERE id =
nexttrunkid;
        -- fill branchid & branchdepth of selected dirrection of other branches
        -- otherbranchids := SELECT ARRAY(SELECT a.e FROM
unnest(nextcrossroad) AS a(e) WHERE a.e <> nexttrunkid);
        otherbranchids := popvalue(nexttrunkid,nextcrossroad);

        UPDATE skeletonmp SET branchdepth = lastbranchdepth + 1 WHERE
ARRAY[id] <@ otherbranchids;
        nextbranchid := maxbranchid;
        IF otherbranchids IS NOT NULL THEN
           FOREACH tmpid IN ARRAY otherbranchids LOOP
              nextbranchid := nextbranchid + 1;
              UPDATE skeletonmp SET branchid = nextbranchid WHERE id = tmpid;
           END LOOP;
                    END IF;
        -- add to array unsolvednodes
        UPDATE skeletonmp SET parrentconnect = 1 WHERE ARRAY[id] <@
nextcrossroad AND ARRAY[curnodeid] <@ touchingendarr;
<@ nextcrossroad;
        IF nextcrossroad IS NOT NULL THEN
           FOREACH tmpid IN ARRAY nextcrossroad LOOP
              SELECT ARRAY[id,parrentconnect] INTO unsolvednodestmp FROM
skeletonmp WHERE id = tmpid;
              IF unsolvednodes IS NULL THEN
                 unsolvednodes := ARRAY[unsolvednodestmp];
              ELSE
                 unsolvednodes := array_cat(unsolvednodes , unsolvednodestmp);
              END IF;
```

```
        END LOOP;
      END IF;
      RETURN unsolvednodes;
         END;
         $$ LANGUAGE plpgsql;
-- test cases a
-- SELECT ARRAY[id,parrentconnect] FROM skeletonmp WHERE ARRAY[id]
<@ ARRAY[48,56];
-- SELECT solvenode(ARRAY[ARRAY[34,0],ARRAY[34,1]]);
-- SELECT solvenode(ARRAY[ARRAY[34,1],ARRAY[36,0]]);
-- SELECT solvenode(ARRAY[ARRAY[36,0],ARRAY[35,0]]);
-- SELECT solvenode(ARRAY[ARRAY[35,0],ARRAY[56,0]]);
-- SELECT solvenode(ARRAY[ARRAY[56,0],ARRAY[38,0]]);
-- SELECT solvenode(ARRAY[ARRAY[38,0],ARRAY[51,1]]);
```

branchWalker.sql

```
DROP FUNCTION branchwalker();
CREATE OR REPLACE FUNCTION branchWalker() RETURNS VOID AS $$
  DECLARE
         dir INT;
         maxdepth INT;
      maxdepthid INT;
      nextcrossroad INT[];
      nextcrossroaddirs INT[];
      lastbranchid INT :=0;
      curbranchazimuth NUMERIC;
      nexttrunkid INT;
      nexttrunkangle NUMERIC;
      unsolvednodes INT[][];
      unsolvednodestmp INT[][];
         BEGIN
         -- create fields
         ALTER TABLE skeletonmp DROP COLUMN IF EXISTS distance;
      ALTER TABLE skeletonmp ADD distance NUMERIC;
      ALTER TABLE skeletonmp DROP COLUMN IF EXISTS azimuth;
      ALTER TABLE skeletonmp ADD azimuth NUMERIC;
      ALTER TABLE skeletonmp DROP COLUMN IF EXISTS turnangle;
      ALTER TABLE skeletonmp ADD turnangle NUMERIC;
      ALTER TABLE skeletonmp DROP COLUMN IF EXISTS turnangleabs;
      ALTER TABLE skeletonmp ADD turnangleabs NUMERIC;
      ALTER TABLE skeletonmp DROP COLUMN IF EXISTS branchid;
      ALTER TABLE skeletonmp ADD branchid INT;
      ALTER TABLE skeletonmp DROP COLUMN IF EXISTS branchdepth;
      ALTER TABLE skeletonmp ADD branchdepth INT;
      ALTER TABLE skeletonmp DROP COLUMN IF EXISTS parrentconnect;
```

```
    ALTER TABLE skeletonmp ADD parrentconnect INT DEFAULT 0;
    -- update distance & azimuth
    UPDATE skeletonmp SET distance = ST_Length(geom);
    UPDATE skeletonmp SET azimuth =
ST_Azimuth(ST_StartPoint(geom),ST_EndPoint(geom));
    -- find core node
    SELECT max(depth) INTO maxdepth FROM skeletonmp;
    SELECT id INTO maxdepthid FROM skeletonmp WHERE depth=maxdepth
LIMIT 1;
    -- fill angles of core node
    UPDATE skeletonmp SET turnangle = 0 WHERE id = maxdepthid;
    -- fill turnangleabs of core node
    UPDATE skeletonmp SET turnangleabs = 0 WHERE id = maxdepthid;
    -- fill branchid & branchdepth of selected dirrection of core node
    UPDATE skeletonmp SET branchid = 0 WHERE id = maxdepthid;
    UPDATE skeletonmp SET branchdepth = lastbranchid WHERE id =
maxdepthid;
    -- get unsolvednodes
    unsolvednodes := ARRAY[ARRAY[maxdepthid,0],ARRAY[maxdepthid,1]];
    -- use solvenode
    WHILE array_length(unsolvednodes,1) >= 1 LOOP
      unsolvednodes := solvenode(unsolvednodes);
    END LOOP;
    -- RETURN unsolvednodes[1][1];
      END;
      $$ LANGUAGE plpgsql;
-- SELECT branchWalker();
```

anaSkel.sql

```
CREATE OR REPLACE FUNCTION anaskel(INTEGER,geometry)
      RETURNS results AS $$
  DECLARE
            turnangleabstest numeric;
      outrow results%ROWTYPE;
      BEGIN
    DROP TABLE IF EXISTS skeleton;
    DROP TABLE IF EXISTS skeletonmp;
    -- CREATE TEMP TABLE skeleton(id SERIAL,geom GEOMETRY);
    CREATE TABLE skeleton(id SERIAL,geom GEOMETRY);
    INSERT INTO skeleton (geom) VALUES ($2);
    -- CREATE TEMP TABLE skeletonmp AS SELECT (ST_Dump(geom)).geom
AS geom FROM skeleton;
```

```
CREATE TABLE skeletonmp AS SELECT (ST_Dump(geom)).geom AS
geom FROM skeleton;
    ALTER TABLE skeletonmp ADD id SERIAL;
    -- PERFORM pgr_nodeNetwork('skeletonmp',0.000001);
            PERFORM updatetouches();
    outrow.id:=$1;
    SELECT count(*) INTO outrow.len FROM skeletonmp;
    PERFORM findDepth();
            PERFORM branchWalker();
    SELECT max(depth) INTO outrow.depthmax FROM skeletonmp;
    SELECT avg(depth) INTO outrow.depthmean FROM skeletonmp;

            SELECT max(touchingn) INTO outrow.touchingnmax FROM
skeletonmp;
    SELECT avg(touchingn) INTO outrow.touchingnmean FROM skeletonmp;

            SELECT min(distance) INTO outrow.distancemin FROM
skeletonmp;
    SELECT max(distance) INTO outrow.distancemax FROM skeletonmp;
            SELECT avg(distance) INTO outrow.distancemean FROM
skeletonmp;
            SELECT sum(distance) INTO outrow.distancesum FROM
skeletonmp;

            SELECT avg(azimuth) INTO outrow.azimuthmean FROM
skeletonmp;

            SELECT min(turnangle) INTO outrow.turnanglemin FROM
skeletonmp;
    SELECT max(turnangle) INTO outrow.turnanglemax FROM skeletonmp;
            SELECT avg(turnangle) INTO outrow.turnanglemean FROM
skeletonmp;
            SELECT sum(turnangle) INTO outrow.turnanglesum FROM
skeletonmp;

            SELECT min(turnangleabs) INTO outrow.turnangleabsmin FROM
skeletonmp;
    SELECT max(turnangleabs) INTO outrow.turnangleabsmax FROM
skeletonmp;
            SELECT avg(turnangleabs) INTO outrow.turnangleabsmean FROM
skeletonmp;
            SELECT sum(turnangleabs) INTO outrow.turnangleabssum FROM
skeletonmp;
```

```
			SELECT sum(turnangleabs) INTO turnangleabstest FROM 
skeletonmp;

			IF turnangleabstest > 0 THEN
				SELECT sum(turnangle) / turnangleabstest INTO 
outrow.wigglines FROM skeletonmp;
			ELSE
				SELECT 0 INTO outrow.wigglinespruned FROM 
skeletonmp;
			END IF;

			SELECT max(branchid) INTO outrow.nrofbranches FROM 
skeletonmp;

	SELECT max(branchdepth) INTO outrow.branchdepthmax FROM 
skeletonmp;
			SELECT avg(branchdepth) INTO outrow.branchdepthmean FROM 
skeletonmp;
	-- prune and repreat var collection
	DELETE FROM skeletonmp WHERE depth = '0';
	PERFORM updatetouches();
	SELECT count(*) INTO outrow.lenpruned FROM skeletonmp;
	PERFORM findDepth();
			PERFORM branchWalker();
	SELECT max(depth) INTO outrow.depthmaxpruned FROM skeletonmp;
	SELECT avg(depth) INTO outrow.depthmeanpruned FROM skeletonmp;

	SELECT max(touchingn) INTO outrow.touchingnmaxpruned FROM 
skeletonmp;
	SELECT avg(touchingn) INTO outrow.touchingnmeanpruned FROM 
skeletonmp;

			SELECT min(distance) INTO outrow.distanceminpruned FROM 
skeletonmp;
	SELECT max(distance) INTO outrow.distancemaxpruned FROM skeletonmp;
			SELECT avg(distance) INTO outrow.distancemeanpruned FROM 
skeletonmp;
			SELECT sum(distance) INTO outrow.distancesumpruned FROM 
skeletonmp;

			SELECT avg(azimuth) INTO outrow.azimuthmeanpruned FROM 
skeletonmp;
```

```
                SELECT min(turnangle) INTO outrow.turnangleminpruned FROM
skeletonmp;
     SELECT max(turnangle) INTO outrow.turnanglemaxpruned FROM
skeletonmp;
                SELECT avg(turnangle) INTO outrow.turnanglemeanpruned
FROM skeletonmp;
                SELECT sum(turnangle) INTO outrow.turnanglesumpruned FROM
skeletonmp;

                SELECT min(turnangleabs) INTO outrow.turnangleabsminpruned
FROM skeletonmp;
     SELECT max(turnangleabs) INTO outrow.turnangleabsmaxpruned FROM
skeletonmp;
                SELECT avg(turnangleabs) INTO outrow.turnangleabsmeanpruned
FROM skeletonmp;
                SELECT sum(turnangleabs) INTO outrow.turnangleabssumpruned
FROM skeletonmp;
                SELECT sum(turnangleabs) INTO turnangleabstest FROM
skeletonmp;

                IF turnangleabstest > 0 THEN
                        SELECT sum(turnangle) / turnangleabstest INTO
outrow.wigglinespruned FROM skeletonmp;
                ELSE
                        SELECT 0 INTO outrow.wigglinespruned FROM
skeletonmp;
                END IF;

                SELECT max(branchid) INTO outrow.nrofbranchespruned FROM
skeletonmp;

                SELECT max(branchdepth) INTO outrow.branchdepthmaxpruned
FROM skeletonmp;
                SELECT avg(branchdepth) INTO outrow.branchdepthmeanpruned
FROM skeletonmp;
     RETURN outrow;
        END;
        $$ LANGUAGE plpgsql;
```

anaSkelWrapper.sql

```
CREATE OR REPLACE FUNCTION anaskelwrapper(TEXT)
        RETURNS VOID AS $$
  DECLARE
        name TEXT := 't' || $1 || 's';
```

```
    BEGIN
        DROP TABLE IF EXISTS skeletons;
        EXECUTE 'CREATE TABLE skeletons AS SELECT * FROM
dblink('
    || quote_literal('dbname=gisdb user=proc password=123')
    || ' , ' || quote_literal('SELECT * FROM '||name)
    || ' ) AS (id integer , geom geometry);';
        DROP TABLE IF EXISTS results0;
        CREATE TABLE results0 AS SELECT anaskel(id,geom) FROM
skeletons;
        DROP TABLE IF EXISTS results1;
        EXECUTE 'CREATE TABLE '
            || quote_ident(name)
            || ' AS SELECT
            (anaskel).id,
            (anaskel).len,
            (anaskel).depthmax,
            (anaskel).depthmean,
            (anaskel).touchingnmax,
            (anaskel).touchingnmean,
            (anaskel).distancemin,
            (anaskel).distancemax,
            (anaskel).distancemean,
            (anaskel).distancesum,
            (anaskel).azimuthmean,
            (anaskel).turnanglemin,
            (anaskel).turnanglemax,
            (anaskel).turnanglemean,
            (anaskel).turnanglesum,
            (anaskel).turnangleabsmin,
            (anaskel).turnangleabsmax,
            (anaskel).turnangleabsmean,
            (anaskel).turnangleabssum,
            (anaskel).wigglines,
            (anaskel).nrofbranches,
            (anaskel).branchdepthmax,
            (anaskel).branchdepthmean,
            (anaskel).lenpruned,
            (anaskel).depthmaxpruned,
            (anaskel).depthmeanpruned,
            (anaskel).touchingnmaxpruned,
            (anaskel).touchingnmeanpruned,
            (anaskel).distanceminpruned,
            (anaskel).distancemaxpruned,
```

```
                        (anaskel).distancemeanpruned,
                        (anaskel).distancesumpruned,
                        (anaskel).azimuthmeanpruned,
                        (anaskel).turnangleminpruned,
                        (anaskel).turnanglemaxpruned,
                        (anaskel).turnanglemeanpruned,
                        (anaskel).turnanglesumpruned,
                        (anaskel).turnangleabsminpruned,
                        (anaskel).turnangleabsmaxpruned,
                        (anaskel).turnangleabsmeanpruned,
                        (anaskel).turnangleabssumpruned,
                        (anaskel).wigglinespruned,
                        (anaskel).nrofbranchespruned,
                        (anaskel).branchdepthmaxpruned,
                        (anaskel).branchdepthmeanpruned
                        FROM results0;';
        END;
        $$ LANGUAGE plpgsql;
-- SELECT anaskelwrapper('51099');
```

Data and code will be available in https://github.com/neuralmax after the dissertation defence.

UDK 711.6(043.3)

SL344. 2021-xx-xx, xx leidyb. apsk. l. Tiražas 16 egz. Užsakymas 263.
Išleido Kauno technologijos universitetas, K. Donelaičio g. 73, 44249 Kaunas
Spausdino leidyklos „Technologija“ spaustuvė, Studentų g. 54, 51424 Kaunas