AN INSTRUMENT FOR COMMUNICATION: THE SELF-ORGANIZING MODEL NIKOLA MARINČIĆ

II

NIKOLA MARINČIĆ is an architect, computer programmer and a musician, working as a postdoctoral researcher at the chair for Computer-Aided Architectural Design at ETH Zurich. He earned a PhD with a medal of distinction from the ETH, with a thesis entitled *Towards communication in CAAD: Spectral Characterization and Modelling with Conjugate Symbolic Domains* in 2017. In 2009 he graduated as an architect from the Faculty of Architecture in Belgrade, and in 2011 completed a Master of Advanced Studies programme in Architecture and Information at ETH Zurich. In 2012–13 he was a guest researcher at the Future Cities Laboratory, interdisciplinary research programme of the Singapore ETH Centre for Global Environmental Sustainability. In 2019, Birkhäuser published his PhD thesis as a part of their Applied Virtuality book series.

INTRODUCTION

The following text is an excerpt from the doctoral thesis entitled *Towards Communication in CAAD: Spectral Characterization and Modelling with Conjugate Symbolic Domains* defended at ETH Zurich in 2017 and awarded with ETH's medal of distinction. The thesis investigates computational models in architecture, a topic usually studied in the field of Computer-Aided Architectural Design (CAAD). Its motivation comes from the author's ever-growing concern with the field's lack of self-reflection and interest to orientate and position itself within the larger context of mathematical and computational modelling. This brings about a general hypothesis that this lack of interest to explore outside of one's field of expertise (often introduced with disciplinarization of computation within a field) might be preventing the field of CAAD from broadening its research domain towards directions that have been successfully explored for a long time in information technology and science. Finally, to show that such a seemingly non-pragmatic effort has an actual research potential, the thesis committed itself to challenge the existing computational models in architecture with contemporary modelling approaches, characteristic for quantum physics, in which computation is regarded from the perspective of *communication* between different domains of a problem. The challenge was actualized in the last part of the work by introducing a new communication-inspired model and applying it in an architectural experiment.

The work engages with an unusually extensive body of knowledge with the aim of providing additional angles to its principal research domain: computational models in architecture. This body of knowledge involves early analytic philosophy, computability and probability theory, formal logic, quantum physics, abstract algebra, computer-aided design, computer graphics, glossematics, machine learning and architecture. However, the reason for such a comprehensive approach and perhaps radical gesture is not to claim any expertise nor mastery over the aforementioned fields of knowledge. To the contrary, it is a matter of methodology, aiming to operate in a more architectural manner, without losing the necessary rigour and consistency required of an academic work.

The first part of the work begins in the 19th century, delves into the body of thinking from which computation emerged and traces two general attitudes towards mathematical modelling, which will each eventually lead to different interpretations of computation. The first one, described as the *logician tradition*, saw the potential of formal, mechanized reasoning in the possibility of constructing the absolute foundation of mathematics, its means of explanation and proof. The second one, the *algebraist tradition*, regarded formalization within a larger scope of model-theoretical procedures, characterized by creatively applying abstraction towards a certain goal. The second attitude proved to be a fertile ground for the redefinition of both mathematics and science, thus paving a way

for contemporary physics and information technology. On the basis of the two traditions, this dissertation identified a discrepancy between the computational models in architecture, following the first tradition, and those commonly used in information technology, following the second.

The Internet revolution, initiated by the development of search engines and social media, is recognized as indicative of the changing role of computers, from ‘computing machinery’ towards the generic infrastructure for communication. In this respect, three contemporary models of communication, proponents of the algebraic tradition, are presented in detail in the second part of the work. As a result, the *self-organizing model* is introduced as the concrete implementation of the ideas appropriated from communication models. This model is described in the following pages.

SELF-ORGANIZING MAP

In a 1976 paper, neurobiologist David Willshaw and physicist Christoph von der Malsburg made a hypothesis that the relation between two laminar structures¹ in the brain is characterized by topographical, system-to-system, rather than cell-to-cell mapping.² Their model was based on a microscopic mechanism that utilized concepts of neural excitation, inhibition and self-organization.³ Inspired by the ideas from neurobiology and neural computation, Finnish physicist and engineer, Teuvo Kohonen generalized the idea of Willshaw and Malsburg in his 1982 paper, *Self-organized formation of topologically correct feature maps*. He stratified his model in two domains and showed that the logic of self-organization is “readily generalizable to mechanisms other than neural.”⁴ To complement the theory, he introduced the self-organizing map algorithm (SOM) aiming to address the question of “how symbolic representations for concepts could be formed automatically.”⁵ His initial discovery was that structure-preserving mapping can be automatically established between input signals in the event space and a topological network responding to those signals.⁶ Steinbuch’s cybernetic notion of a learning matrix (Lernmatrix) inspired Kohonen to define the response layer of his model as a memory matrix.⁷

1 D. J. Willshaw, and C. V. D. Malsburg, ‘How Patterned Neural Connections Can Be Set Up by Self-Organization’, *Proceedings of the Royal Society B*, 194, no. 1117 (November 12, 1976): 431–45, (433). Sheets of brain cells, namely pre- and postsynaptic.

2 Ibid., 442.

3 Ibid., 433.

4 Teuvo Kohonen, ‘Self-Organized Formation of Topologically Correct Feature Maps’, *Biological Cybernetics*, 43, no. 1 (1982): 59–69, (59).

5 Ibid., 59.

6 Ibid., 59.

7 “Another early construct which was introduced for the explanation of brain functions and implementation of artificial intelligence is the Learning Matrix due to Steinbuch. It is a system of crossing signal lines, with an adaptive connection at each crossing.” Teuvo Kohonen, *Self-Organization and Associative Memory*, 2nd ed. Springer Series in Information Sciences, eds. T. Huang and M. Schroeder, vol 8 (Berlin: Springer, 2012). 74.

FIG 1
Kohonen's illustration
of a system
which implements
an ordered mapping.

SOM AS A REPRESENTATION

Kohonen showed that different network topologies could be implemented as a response layer of his model.⁸ The topology of a two-dimensional lattice revealed some advantages in this respect: It yielded good results, allowed for easy visualization⁹, and complied with the idea of brain maps.¹⁰ The output domain of Kohonen's algorithm was named a *map*, and the algorithm the *self-organizing map*.¹¹ The naming also implied that the communication between different domains was in fact a *mapping*. The two-dimensional representation of the self-organizing map algorithm influenced its interpretation in the artificial intelligence circles.¹² In the mid 1990s, the availability, memory and speed of computers increased enough to accommodate for the high computational cost of the algorithm. Researchers attained the means to conduct more experiments and demonstrated novel capabilities of the algorithm, many of which were completely unrelated to visualization and analysis.¹³ Thirty years later, Kohonen acknowledges a surprisingly wide variety of applications and modifications of the algorithm and tries to

8 Teuvo Kohonen, 'Essentials of the Self-Organizing Map', *Neural Networks*, 37 (January 2013): 52–65 (55–56)

9 "... the 2-D grid allows rough visual presentation and interpretation of the clusters." Vesanto and Alhoniemi, "Clustering of the Self-Organizing Map," 599.

10 "It is believed that many other kinds of maps exist in the hippocampus or other parts of the brain system." Kohonen, *Self-Organization and Associative Memory*, 121; ... for example *tonotopic maps* in auditory cortex of the brain, *Ibid.*, 120.

11 "The earliest SOM models, tending to replicate the detailed neural-network structures, were intended for the description and explanation of the creation of brain maps." Kohonen, "Essentials of the Self-Organizing Map," 55.

12 Elaborated in the next sub-chapter.

13 Such as: Guilherme Barreto and A. Araújo, 'Identification and Control of Dynamical Systems using the Self-Organizing Map', *IEEE Transactions on Neural Networks*, 15, no. 5 (September 1, 2004): 1244–59, and Guilherme Barreto and L. Souza, 'Adaptive Fitering with the Self-Organizing Map: A Performance Comparison', *Neural Networks*, 19, no. 6 (July 2006): 785–98.

were visualization and data analysis.¹⁸ In a similar manner, Vesanto concurred that the algorithm has shown itself valuable in many applications, but most prominently in visualizing high-dimensional spaces.¹⁹ In one of his early papers, he made an overview of different visualization methods.²⁰ A year later, Vesanto employed a k-means algorithm on top of the SOM response layer and concluded that in many cases SOM clustering yields better results than other clustering techniques.²¹ However, he found this to be a fortunate side-effect of the SOM's powerful visualization abilities.

FIG 3
Example of Ultsch's
U*Matrix of the
'Glio' data set.

- ¹⁸ "Accordingly, the most important applications of the SOM are in the visualization of high-dimensional systems and processes and discovery of categories and abstractions from raw data. The latter operation is called the exploratory data analysis or *data mining*." Kohonen, Teuvo, et al., 'Engineering Applications of the Self-Organizing Map', *Proceedings of the IEEE*, 84, no. 10 (October 10, 1996): 1358–83 (1358).
- ¹⁹ "Visualization of complex multidimensional data is indeed one of the main application areas of the SOM." Juha Vesanto, 'SOM-Based Data Visualization Methods', *Intelligent Data Analysis*, 3, no. 2 (January 1, 1999): 111–26 (112).
- ²⁰ *Ibid.*, 115–24.
- ²¹ "In [16], partitive methods (i.e., a small SOM) greatly outperformed hierarchical methods in clustering imperfect data." Vesanto and Alhoniemi, 'Clustering of the Self-Organizing Map', 588.
- "Another benefit is noise reduction. The prototypes are local averages of the data and, therefore, less sensitive to random variations than the original data." *Ibid.*, 588.
- "The experiments indicated that clustering the SOM instead of directly clustering the data is computationally effective approach." *Ibid.*, 599.

In order to reach a better understanding of the capabilities of self-organizing maps, work by other researchers focused on the applications of the algorithm reaching beyond mere analysis and representation,²² including nonlinear time series prediction, function approximation, control of dynamical systems²³, nonlinear adaptive filtering²⁴, data exploration (in industry, finance, natural sciences, and linguistics), digital signal processing/transmission²⁵, and nonparametric encoding for high-dimensional spatiotemporal dynamics.²⁶ In his doctoral thesis on pre-specific modeling, Moosavi provided a rich list of indexes around the SOM algorithm.²⁷

SELF-ORGANIZING MODEL

If the dominant theoretical interpretations of the algorithm are correlated to the diversity of its applications indexed by Moosavi, there appears to be an unsettling discrepancy between the platitude of interpretations and the actual demonstrated capacity. How can a tool for visualization or an automatic data-analysis method be so successfully used for function-approximation, control of dynamical systems, or prediction of time-series? The important hypothesis of this work is that the insistence on the terms *map* and *mapping* to describe Kohonen's algorithm reduces it to a visualization and analytical tool, hence obscuring the actual capacity demonstrated in applications. Such interpretations systematically neglect higher-level architectonics on top of which the algorithm is conceived, thus, in our opinion, invoke erroneous theoretical consequences. The ambition of this work, motivated by the work of Hovestadt, Bühlmann, Moosavi, Wassermann, Zafiris, and Hjelmlev²⁸, is to propose an alternative

22 " ... However, most of this methods do not use the representational power of the SOM to build the prediction models." Barreto and Araújo, "Identification and Control of Dynamical Systems Using the Self-Organizing Map," 1257.

23 Barreto and Araújo show that "the SOM can be successfully used to approximate dynamical input-output mappings, with minor modifications in the original algorithm." Ibid., 1244.

24 For example: identification and equalization of communication channels, Barreto and Souza, "Adaptive Filtering with the Self-Organizing Map: a Performance Comparison," 787.

25 Kohonen, 'Essentials of the Self-Organizing Map', 52.

26 Vahid Moosavi, 'Computing with Contextual Numbers', *ArXiv abs/1408.0889* (August 5, 2014).

27 Vahid Moosavi, 'Pre-Specific Modeling: Computational Machines in a Coexistence with Concrete Universals and Data Streams', in *Metalithikum IV, Coding as Literacy*, ed. by Vera Bühlmann, Ludger Hovestadt, and Vahid Moosavi (Vienna: Birkhäuser, 2015), 132–66 (71).

28 see Birkhäuser's Applied Virtuality Book Series edited by Hovestadt and Bühlmann; replace with: Elias Zafiris, *Natural Communication: A Functorial Approach*. Applied Virtuality Book Series, edited by Ludger Hovestadt and Vera Bühlmann. (Basel: Birkhäuser, forthcoming). Moosavi, 'Pre-specific Modeling'; *The Putnam Program: Language & Brains, Machines and Minds* accessible at: <https://theputnamprogram.wordpress.com>.

theoretical interpretation of the self-organizing map from a more abstract mathematical perspective. The hypothesis is that the very architectonics inherent to the algorithm, not only its specific implementations, makes it unique among the computational frameworks. Such interpretation should provide an adequate account for its seemingly extraordinary capacities.

REASON AND COEXISTENCE

The idea of a system is not new, but its usual interpretation within the computer science circles is usually credited to Norbert Wiener. This is not very surprising as computer science “directly applies the concepts of cybernetics to the control of devices and the analysis of information,” such as cellular automation, decision support, robotics and simulation.²⁹ Cybernetics had a crucial influence on artificial intelligence, computer modelling and artificial neural networks, under which the SOM is usually classified.³⁰ Hovestadt argues how the cybernetic notion of a system is conceived upon the notions of *reason* and *causality* (in a Newtonian sense), which makes it incompatible with today’s body of thinking conceived upon the invention of electricity and the development of quantum physics.³¹ Notions of reason and a cause are also constituents of the scientific method. In his prolegomena, Hjelmslev described the role of these notions in defining the limits of scientific knowledge:

Science always seeks to comprehend objects as consequences of a reason or as effects of a cause. But if the object can be resolved only into objects that may all indifferently be said to be consequences or effects of all or none, a continued scientific analysis becomes fruitless.³²

The book, *Coding as Literacy*,³³ appearing in Birkhäuser’s *Applied Virtuality Book Series*, discussed the potential of Self-Organizing Maps as a model³⁴ from the perspective of computational literacy. In the chapter on elements of digital architecture, Hovestadt takes a quantum perspective to invert the customary interpretation of the SOM as a map. He gave an example of the self-organized map of Zurich where

29 Wikipedia, *Cybernetics*, last modified August 30, 2017, 17:55, <https://en.wikipedia.org/wiki/Cybernetics>.

30 Wikipedia, *Cybernetics*.

31 As an overview of Hovestadt’s position on the topic see: Ludger Hovestadt, ‘Elements of Digital Architecture’, in *Metalithikum IV, Coding as Literacy*, ed. by Vera Bühlmann, Ludger Hovestadt, and Vahid Moosavi (Vienna: Birkhäuser, 2015), 28–114. And Ludger Hovestadt, ‘Cultivating the Generic: A Mathematically Inspired Pathway for Architects’, in *EigenArchitecture: Computability as Literacy*, ed. by Ludger Hovestadt and Vera Bühlmann (Vienna: Ambra Verlag, 2014), 7–67.

32 Louis Hjelmslev, *Prolegomena to a Theory of Language*, trans. F. J. Whitfield (Wisconsin: The University of Wisconsin Press, 1963), 84.

33 Vera Bühlmann, Ludger Hovestadt, and Vahid Moosavi, eds., *Coding as Literacy: Metalithikum IV*, Applied Virtuality Book Series (Basel: Birkhäuser, 2015).

34 “The self-organizing map articulates the logical form of chronological elements in probability values ... Therefore, we suggest that we should not talk about a self-organizing map but a self-organizing model.” Hovestadt, ‘Elements of Digital Architecture’, 108.

“the constellation of elements changes according to the analytical/chronological position of the observer.”³⁵ In his doctoral thesis, Moosavi further investigated the notion of coexistence in the context of urban data streams and showed the scope of applicability of data-driven models.³⁶ Inspired by the ideas of Wassermann³⁷, Hovestadt and Bühlmann, this work reinterprets self-organizing maps by identifying the role of the algorithm within a more general, *self-organizing model*. This model abandons the idea of systematicity based on reasoning and facilitates the notion of *coexistence* in the world of data.

KOHONEN’S ALGORITHM AS A KEY

If we put the emphasis on the implementation of Kohonen’s algorithm, we can acknowledge the existence of two data structures resembling different structural domains.³⁸ Kohonen’s input layer is implemented as an n -dimensional list of m -dimensional vectors, where n and m are positive integers. The response layer (or map³⁹) is implemented as a two-dimensional array (the common encoding of a graph/matrix) of m -dimensional vectors, where the dimensionality m of the input layer vectors, matches the dimensionality of the response-layer vectors.⁴⁰ Kohonen’s crucial discovery demonstrated by the algorithm is that the input signal alone is sufficient to enforce self-organization in the response layer, but only if the response layer has a specific architecture and implements a local feedback mechanism.⁴¹ The response layer (map) should then be interpreted by acknowledging that the mechanism imposed on the set of input vectors forces unlabelled classes to emerge in the response layer, with the effect that the topological distances within the response layer resemble similarities between the classes. The interesting discovery is that the schema of Kohonen’s algorithm can be accurately represented by means of a key of the same anatomy as Zafiris’s motivic key. Here, the input and output layers could be interpreted as different domains, and mapping and interpretation as adjoint communication bridges.

35 Ibid., 109.

36 If we assume “... each of the prototypical forms in the trained SOM metaphorically as a word or a letter in a language, a trained SOM in coexistence with data can be interpreted as a pre-specific dictionary of the target phenomena.” Moosavi, ‘Pre-Specific Modeling’, 73.

37 See: *The Putnam Program*.

38 “This principle is a generalization of the formation of direct topographic projections between two laminar structures known as retinotectal mapping.” Kohonen, ‘Self-Organized Formation of Topologically Correct Feature Maps’, 59.

39 ... also known as lattice in case of 2-dimensional response-layer.

40 Ibid., 61.

41 Ibid., 59.

FIG 4
SOM algorithm represented in a form of Zafiris's motivic key

On the level of technical implementation, the elements of this schema appear to be too specific if compared with the Zafiris's motivic key. However, this is an important first step towards building a more abstract perspective. As a next step, I will demonstrate that successful SOM applications are in fact taking advantage of inverting or reinventing one or more elements of key defined as such.

MODIFICATIONS OF THE INPUT LAYER

In their paper *Neural methods for non-standard data*, Hammer and Jain explored the possibilities of different structures serving as an input layer of the algorithm. Instead of vectors, they explored the possibility of utilizing strings, sequences, trees, graphs or functions.⁴² In his paper *Computing with contextual numbers* Moosavi created a two-level input layer structure by replacing the original layer with a one-dimensional self-organizing map, acting as a non-parametric encoder.⁴³ These approaches are very powerful as they allow much greater freedom in the constitution of the encoding.

MODIFICATIONS OF THE RESPONSE LAYER

The work of Alfred Ultsch involves many experiments on the output layer of the self-organizing map. To minimize the learning error, he explored different network topologies in the response layer, for example hexagonal and rectangular grids.⁴⁴ U-matrix, U*-matrix and U*C-matrix algorithms directly exploited the response layer of the SOM. They handled the response layer, its distances and densities, as an input for clustering. Algorithms like neural gas and growing neural gas, replaced the two-dimensional lattice with a flexible and adaptive response layer.

42 Barbara Hammer and B. Jain, 'Neural Methods for Non-Standard Data', *ESANN 2004 Proceedings*, 281–292, Bruges, Belgium (2004), 289.

43 Moosavi, 'Computing with Contextual Numbers', 13.

44 Alfred Ultsch and L. Hermann, 'The Architecture of Emergent Self-Organizing Maps to Reduce Projection Errors', *Proceedings of the European Symposium on Artificial Neural Networks (ESANN)*, 1–16, Bruges, Belgium (2005), 2.

Such a response layer could subsume any topology as it dynamically adapted itself to the variability of the input signals.⁴⁵

MODIFICATIONS OF THE MAPPING/LEARNING

The generative topographic mapping algorithm (GTM) was a counterpart to the SOM, which tried to overcome its limitations by rethinking the *mapping* between the domains. For that purpose, it uses a “constrained mixture of Gaussians whose parameters can be optimized using the EM (expectation-maximization) algorithm.”⁴⁶

MODIFICATIONS OF THE INTERPRETATION

Without attempting any interpretation of the results of SOM training, Barreto and Souza utilized the nodes of the response layer as the reference inventory for prediction by using the *k*-nearest neighbours algorithm.⁴⁷ The vector, whose missing value (dimension) was to be predicted was first introduced into the response layer and compared to those of its vectors without the missing value. The influence of every node was calculated according to their relative distance and the missing value interpolated. For the control of dynamical systems, Barreto and Araújo established a feedback loop between the input and response layer, thus enforcing the dynamic control of a system.⁴⁸

COEXISTENCE KEY

The intention behind presenting these applications was to demonstrate that the modification or replacement of the particular elements of SOM architecture need not change the form of the key used to describe its architecture. In this regard, what differentiates SOM from other machine learning algorithms lies in its structural properties. Hjelmlev's differentiation between variants and invariants can help us to understand why this might be the case. His conception and terminology support the argument that all the presented modifications of the original algorithm are in fact *variants* by the virtue of placing a specific algorithmic procedure to a specific position

45 Bernd Fritzke, 'A Growing Neural Gas Network Learns Topologies', *Advances in Neural Information Processing Systems*, 7 (1995), 625–32 (626).

46 “... lack of a theoretical basis for choosing learning rate parameter schedules and neighbourhood parameters to ensure topographic ordering, the absence of any general proofs of convergence, and the fact that the model does not define a probability density.” Christopher Bishop, M. Svensén and C. Williams, 'GTM: The Generative Topological Mapping', *Neural Computation*, 10, no. 1 (January 1998), 215–34 (216).

47 Guilherme Barreto, João C. M. Mota, Luis G. M. Souza, and Rewbenio A. Frota, 'Nonstationary Time Series Prediction Using Local Models Based on Competitive Neural Networks', in *Innovations in Applied Artificial Intelligence*, ed. by Bob Orchard, Chunsheng Yang, and Moonis Ali (Berlin, Heidelberg: Springer Berlin Heidelberg, 2004), 1146–55 (1146).

48 Barreto and Araújo, 'Identification and Control of Dynamical Systems Using the Self-Organizing Map', 1245.

within the key. Consequently, the *invariant* property of the self-organizing model is reflected in the position of an element within the key, and is defined relationally in respect to other elements. This is an indicator that there is more to the SOM than just an algorithm. It is a computational framework resting upon a mathematical contract established by the motivic key and its invariant anatomy. We can define such a key as a *coexistence key* and such a framework as the *self-organizing model*. The coexistence key maintains a contract established upon mathematical symmetries, while the self-organizing model furnishes its elements with computational procedures.

FIG 5
Coexistence key of
the self-organizing
model.

In the manner of Zafiris’s motivic key, the coexistence key stratifies the self-organizing model in two levels, as shown on Figure 5. The domain *A* is defined as an unordered set of observables, and domain *B* as a self-referential system⁴⁹ of objects and arrows. According to Kohonen, the input signal alone is sufficient to enforce self-organization in the response layer, if the response layer has a ‘specific architecture’ and implements a local feedback mechanism.⁵⁰ The proposed stratification gives new insights to Kohonen’s requirement. If we are interested in what would be the invariant property of Kohonen’s ‘specific architecture’—concerning we have acknowledged that many different structures could successfully serve as the response layer—an answer comes from the category theory. A domain which allows to be self-organized must be self-referential. As we have seen with growing neural gas algorithm, any topological structure allows for self-organization as long as every node has a reference to itself and to all the other nodes.⁵¹ At a computational colloquy held at the CAAD chair, Hovestadt presented a series of experiments on the coexistence between data and the wide range of self-referential topological structures, ranging from linear chains to multiple graphs. To represent

49 The notion of the system is taken from Hjelmlev, meaning premised system. Hjelmlev, *Prolegomena to a Theory of Language*, 16.

50 Kohonen, ‘Self-Organized Formation of Topologically Correct Feature Maps’, 59.

51 Ibid., 60, fig. 2.

the model, I suggest using the category theory instead of graph theory, as it is better suited for representing its structural properties.

Metaphorically speaking, domains *A* and *B* can be regarded as different ‘natures’ of things.⁵² From the perspective of Hjelmslev’s glossematics, an obstacle or a problem observed in the domain *A* would be a specific purport formation (substance), apparent to us by the virtue of its content-form encatalysed within the same domain. By introducing another domain, we are seeking to find a new purport formation within the domain *B* describing the same obstacle or a problem by the virtue of a different content-form.

In order to establish coexistence between the domains *A* and *B*, the model uses encoding and decoding bridges. For the encoding bridge, the self-organizing model requires a generalization procedure that can produce a partition spectrum from an indiscrete, unordered set of observables. The self-organizing map algorithm provides that functionality. In the process of self-organization, the pockets of negentropy emerge from the entropic data set resulting with a partition spectrum.⁵³ Once the spectrum is obtained, the obstacle or a problem is represented by means of a new cipher (a category) within the spectral domain *B*. Let us assume that in the domain *B* there exists a procedure allowing to bypass the obstacle, or solve the problem. In order to interpret what was resolved in the spectral domain *B*, the self-organizing model uses a decoding bridge. The decoding bridge requires a projection procedure that utilizes virtual distinctions from the partition spectrum to distribute the original set of observables (encoded as the unordered set) into a category of sets corresponding to actual distinctions. The actual distinctions give the interpretation to the solution/obstacle within the original domain.

FORM AND SUBSTANCE: ENCODING THE BLACK BOX

According to Zafiris, the notion of data assumes a geometrical encoding. Hjelmslev sees data as a *content-substance*—formed purport encatalysed by means of the content-form.⁵⁴ This suggests that the encoding principle itself is a *form*. In a chapter from *A Thousand Plateaus*, Gilles Deleuze has multiplied this elementary gesture until he stratified “the whole earth” and beyond, by means of coding and territorialization.⁵⁵ This gesture introduces a very important modelling question: What is the role of a form

52 This is in the sense that physical laws that apply in the quantum scale, seem to be completely different from those that apply in the macroscopic scale, and according to that they could be considered as different “natures” of things.

53 Resolution of this spectrum is directly proportional to the number of nodes (objects) that were selected to constitute the network, and its number can be larger or smaller than the number of elements in the initial set of observables.

54 Hjelmslev, *Prolegomena to a Theory of Language*, 54.

55 Gilles Deleuze and F. Guattari, ‘10,000 B.C.: The Geology of Morals; Who Does the Earth Think It Is?’, in *A Thousand Plateaus*, trans. B. Massumi (London: Bloomsbury Academic, 2004), 44–82 (45–48).

in the context of a stratified model? Zafiris gives an example of applying his model to quantum physics, where the motivic key enables communication between two physical levels: atomic and microscopic. On an atomic level, no technical means are sufficient to locate an electron and its interactions. Thus, the whole domain can be thought of as a black box. On a microscopic level, physicists observe interference patterns and think of them as a reflection of unknown events happening on the atomic level. For Zafiris, modelling consists of establishing a pair of functors from the observational universe to the microscopic universe and back.⁵⁶ In this way, we can regard these two levels as functors, one of which is giving form to another. In this way, a form is given to something that cannot be observed.

This perspective must be investigated from the standpoint of the coexistence key. By applying the same gesture, we can consider the domain *A* of the coexistence key as a 'black box' characterized by a set of observables. We can assume that the observables describe some meaningful process, but the semantics is not explicitly present in the observables. Starting from the domain *A*, the application of the generalization bridge will establish a partition spectrum. The resulting spectrum is constructed in another cipher, and at the same time juxtaposed with our original set of observables (a black box). The new cipher of domain *B*, known as a *category*, comes with a greater operative and symbolic potential than the cipher of the original domain (set).

However, we are still missing the process of formalization. In order to give form to the domain *B*, the system of objects and arrows must be composed in such a way that it resembles relations between the observables in the domain *A*. The role of the generalization bridge is to give form to the domain *A* within the domain *B*. In the same way, we can see the projection bridge as responsible for obtaining semantics in the domain *A* according to the form of the domain *B*. By doing so, the projection bridge closes the path and establishes coexistence between levels *A* and *B*.

PROCEDURAL CATALYSIS OF SYNCRETISMS

In order to formally define the process of formalization, Hjelmslev introduced the term catalysis. The term implies that "form is encatalyzed to substance and language encatalyzed to the text."⁵⁷ If we apply the concept of catalysis to the self-organizing model, instead of saying that the domain *B* gives form to the domain *A*, we can say that the self-organization procedure encatalyzes the form of the category to the set of observables. If the first criterion that makes self-organization possible is self-reference, the second criterion would be the possibility of

⁵⁶ Zafiris, *Natural Communication*.

⁵⁷ Hjelmslev, *Prolegomena to a Theory of Language*, 96.

information exchange between the nodes of the system. Such property, Kohonen calls a local feedback.⁵⁸ In the self-organization procedure, objects within the system are pushed/pulled towards each other as they exchange information. The mechanism which enforces this behaviour in the SOM algorithm is known as the distance function.⁵⁹ This function provides a metrical criterion of similarity. By applying the concept of catalysis, we can say that the distance function locally encatalyzes the form to the set of affected observables. Influenced by the distance function, nodes receiving the local feedback begin to share the same-expression form. This is especially visible between two directly connected nodes of the system.

Such behaviour corresponds very well to the Hjelmslev's definition of syncretism.⁶⁰ It is a category established upon the suspended mutation which he called overlapping.⁶¹ Therefore, we can regard the distance function as the form which is encatalyzed locally to the set of observables and establishes overlapping between the nodes within the system. In other words, the nodes are organized in such a way that they will be 'pushed' to share the same expression-form according to the applied distance function. Hjelmslev uses the notion of syncretism to define the term concept:

concept—a syncretism between things (namely, the things that the concept subsumes).⁶²

According to the definition, and by the virtue of the distance function, we can define the population of nodes that exhibit overlapping, a *concept*. This can be illustrated with an example of an experiment that the author conducted in 2012. Here, the self-organizing map algorithm was used to explore design possibilities of a generic parametric model. The model consisted of six parameters that defined the shape and colour of a single geometric figure. This figure was replicated in one thousand instances with random parameter values and the SOM algorithm was used to organize them into a two-dimensional lattice. As Figure 6 shows, the outlined collection gathers locally similar shapes, but as we 'zoom out' and start to consider the population of shapes, we can notice the difference between shapes belonging to different collections. From the standpoint of expression-form, each two neighbours within a single collection are overlapping. However, if we select two shapes belonging to different collections, they are most likely not overlapping. Those that overlap, can be said to establish a category defined by the overlapping. Thus, their collection, as the one marked on the image, can be regarded as the concept.

58 Kohonen, 'Self-Organized Formation of Topologically Correct Feature Maps', 59.

59 Kohonen, 'Essentials of the Self-Organizing Map', 53.

60 "... the category that is established by an overlapping we call (in both planes of a language) a syncretism." Hjelmslev, *Prolegomena to a Theory of Language*, 88.

61 Ibid., 88.

62 Ibid., 93.

FIG 6
Interpretation of
Hjelmslev's notion
of a concept within a
partition spectrum.

From the perspective of glossematics, the relational property of the model that enforces that each individual shape locally reflects the arrangement of all the others, we can call a *connotator*. Therefore, in the previous example, each individual shape is considered to be the connotator of the category it belongs to.

EXPERIMENT

In the last part of the doctoral thesis, the self-organizing model is applied in an architectural experiment. Its aim is to address the question of how similarity between spaces in architecture could be accounted for within a paradigm of communication. The hypothesis is that the question of similarity cannot be adequately posed within the established computational models in architecture, due to their reliance on a pre-given notion of space in which the architectural representations are submerged. Being able to speak about the similarity of spaces requires addressing the problem outside of the given space of representation, finding or constructing another level of representation of the same phenomena, where the question can be posed and answered once the coexistence between two levels is established. By applying the self-organizing model, the author attempts to show that this is indeed possible. The experiment was defined as following:

The objects of modelling in this experiment are floor plans. After addressing the questions of conventions of floor plan representation, its scale, format, and availability in terms of data, the decision had been made to operate on raster images of floor plans.

The 2009 book *typologie+*⁶⁵ contained 189 floor plans that satisfied the requirements of the experiment and that were digitally scanned, cropped in high resolution at the same scale and stored in raster format. The plans' computational ambiguities were resolved in the data pre-processing stage, namely representation of walls' interior as empty space bounded by lines, and deciding how the plans should be oriented (Figure 7).

63 Ebner, Peter, Eva Herrmann, Roman Höllbacher, Markus Kuntscher, and Ulrike Wietzorrek. *typologie+*: *Innovativer Wohnungsbau* (Basel: Birkhäuser, 2009).

FIG 7
The final data
set after the
pre-processing stage.

At this stage, it was possible to think of the collection of floor plans as *an abstract object* characterized by the dataset of raster images. The codes and ciphers found at this level (graphical and symbolic elements and their conventions) are familiar to every architect, but do not provide an adequate basis for the comparison of floor plans. The aim is to establish a completely new level, where the same abstract object can be represented, but in entirely different terms. The codes and ciphers on this new level are expected to provide a more adequate basis for comparison. For this reason, the choice was made to construct the new level on a symbolic basis. The first and most important step towards this goal is the creation of communication bridges.

The modelling begins by partitioning this object in order to obtain the partition spectrum. The partition of each individual floor plan into parts is accomplished by dividing the image in such a way, that the collection of equal size parts covers the image entirely, with a variable amount of overlapping between the parts.

When the partition procedure was applied to all 188 floor plan images with a desired overlapping, and after applying a rotation transformation group to each part, the abstract floor plan object was represented as a collection of 573,488 image parts. Figure 8 shows around 23,000 image parts extracted from it.

FIG 8
23,393 floor plan
parts on a single
image

What follows is an application of the translation and mirroring invariance bridge that translates each floor plan part into a histogram of frequencies representing the lengths of stretches of black pixels (obstacles) registered in scanning each part in horizontal and vertical direction line by line. This procedure helped to bypass a complicated problem of how to appropriately choose the position of each individual cut.

In the next stage, the generalization bridge of the self-organizing model is applied to the collection of elements (each now represented as a list of 20 characteristic frequencies) which yields a partition spectrum. Kohonen's self-organizing map algorithm serves as the implementation of this generalization bridge. The first level of the self-organizing model was defined as an unordered set of observables represented by the collection of floor plan parts represented as 20-dimensional frequency vectors. The generalization procedure connects this level to another level—a category—represented by a self-referential system of objects and arrows, with objects have the same dimensionality, but different connections.

By applying partition and generalization procedures of the self-organizing model to a large number of floor plan images, a finite collection of elementary geometric expressions was extracted, and a symbol attached to each instance (Figure 9).

FIG 9
10,000 randomly
chosen floor plan
parts extracted from
the initial data set
and clustered into
100 sets of actual
distinctions.

This collection of symbols is regarded as the alphabet and any floor plan created by the same conventions, can thereby be defined on this symbolic basis. Finally, each floor plan is represented as a chain of probabilities, based upon its individual alphabetic expression of a written language, and its values used to compute similarities between plans. The Figure 10 shows the space of floorplan similarities, where the similarity of any two plans *A* and *B* taken from the original data set can be expressed by the topological distance between their respective cells.

FIG 10
Topological distance
between the cells as
the measure of floor
plan similarity.

EPILOGUE

The experiment attempted to set the self-organizing model in motion by implementing it in an architectural context. It was demonstrated how a contemporary concept of communication could be applied to architectural modelling. Despite the real benefits of the symbolic representation applied to floor plans, this latter may prove to be insufficient in future research. This would miss the point of the experiment. Its object might be specific, but its aim is not: to present a novel perspective on how the concept of similarity could be derived from an object and in an object’s own terms, and, furthermore, how we can computationally encode such an idea by learning from seemingly very different fields such as quantum physics, not the most obvious research context for architecture. One of the promises of communication models such as the self-organizing model is that they allow us to overcome the limiting assumptions of set theory about the constitution of the very modelling. Objects we are trying to model need no longer be sharply defined or have clear boundaries in order for us to encode them computationally. They may be hard, but also soft, or any state in between. This is due to their fluid identity which oscillates as we change the ratio between the information and symmetry we are willing to extract or maintain in each modelling scenario. As a consequence, communication models allow us to take modelling seriously and rigorously, as it is expected in a scientific context, while drastically extending the amount of contingency of what can be articulated in favour of old necessities.