COMMENTARY

CGGC: Category-based General GeoComputation

Linwang Yuan,[1,2,*] Zhaoyuan Yu,[1,2] Wen Luo,[1,2] Dongshuang Li,[2,3] Jiyi Zhang,[4] Liangchen Zhou,[1,2] and Guonian Lü[1,2]
[1]Key Laboratory of Virtual Geographic Environment (Ministry of Education of PRC), Nanjing Normal University, Nanjing 210023, China
[2]Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China
[3]School of Marine Science and Engineering, Nanjing Normal University, Nanjing 210023, China
[4]School of Geographical Science, Nantong University, Nantong 226019, China
*Correspondence: yuanlinwang@njnu.edu.cn
Received: June 5, 2024; Accepted: November 7, 2024; Published Online: November 26, 2024; https://doi.org/10.59717/j.xinn-geo.2024.100102

Citation: Yuan L., Yu Z., Luo W., et al., (2024). CGGC: Category-based General GeoComputation. The Innovation Geoscience **2(4)**: 100102.

GeoComputation occupies a strategic position as the bedrock of smart cities and government operations. Traditionally, GeoComputation has been grounded in reductionism, where intricate geographic systems are deconstructed into discrete elements for comprehension and analysis. However, as the scope of geographic information applications widens, GeoComputation requirements have become more complex, necessitating intricate computational models and structures across diverse domains. There is an urgent need to build a comprehensive, top-level general GeoComputation frameworks, and to facilitate adaptable linkages between cross-domain data representation, processing mechanisms, and computational services.[1]

GeoComputation has evolved through three distinct epochs: mainframe computing, personal computing, and ubiquitous computing,[2] corresponding to integrated, component-based, and domain-oriented design patterns (Figure 1). Of the three patterns, domain-oriented GeoComputation is the closest cross-domain solution. Coupled with the advent of the Internet and cloud computing, it paved the way for microservices-driven development, however, from the reductionist and set perspective, it still encounters fragmentation in computational constructs. In addressing multi-domain analytical challenges within smart cities and digital societies, general GeoComputation is found to be lacking. A holistic perspective is needed, along with the development of flexible, generative, and adaptable approaches.

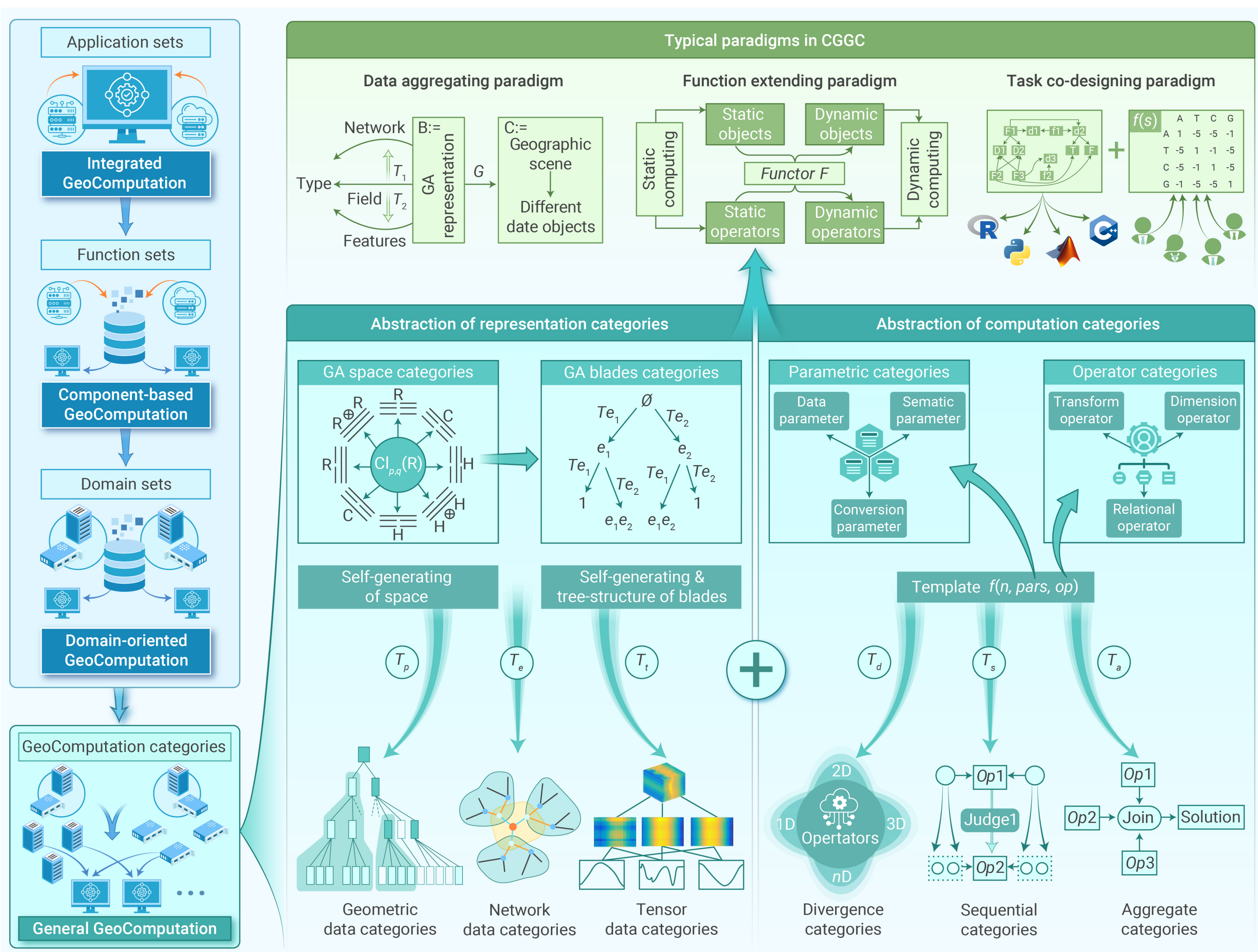

Figure 1. The framework of Category-based General GeoComputation (CGGC)

General GeoComputation needs to establish a universal framework across diverse domains to achieve semantic harmony between the expression architecture and computational architecture of participating entities. Category theory, which transcends set theory as a universal abstraction, describes the mathematical theory of general abstract structures through objectoriented and mapping-oriented perspectives.[3-4] Category theory formalizes the algebraic structure concept into a series of "objects" and "morphisms", studies the isomorphism of "morphisms" in an axiomatized manner, and constructs "functors" to associate mathematical structures. The mitigation of differences between objects enables the synthesis and extension of mapping patterns, abstracting universal relationships through composition. The high extensibility of Category Theory has led to its wide use in software engineering, system control, graph theory, natural language understanding and many other fields.

This paper proposed the Category-based General GeoComputation (CGGC) model. Unlike set theory, category theory offers a more abstract viewpoint that emphasizes the essence of operations rather than individual elements. It allows for the precise description of commonalities across diverse fields in an abstract manner. Furthermore, category theory simultaneously captures mappings between objects, categories, and relationships to integrated various types of data, operators, and domain structures. Through these structured associations, domains and relationships can be combined to form new domains or connections, enabling a holistic approach to GeoComputation.

CONSTRUCTING OF REPRESENTATION AND COMPUTATION CATEGORIES

CGGC builds upon the systematic organization and abstraction of application sets, function sets, and domain sets in reductionist GeoComputation. Initially, data categories, function categories, and domain categories are constructed to form the basic objects of CGGC. Morphisms are then used to connect these objects, while functors and natural transformations facilitate mappings and transformations between these three categories. Although CGGC possesses significant advantages, its limited adoption in the GIS field can be attributed to the lack of a formal GIS representation and computation framework, which hinders the categorical expression and computation of GeoComputations. The proposed multidimensional unified data model, grounded in Geometric Algebra (GA), represents an innovative endeavor at the foundational level of mathematical theory,[5] which provides a foundation for constructing representation and computation categories for geographic information (Figure 1).

Abstraction of representation categories

GA is a powerful tool that achieves generality in spatial representation by unifying various algebraic spaces. It accomplishes this by utilizing positive, negative, and null spaces, thereby creating a self-generated GA space category. Using the well-defined outer product, blades can be generated as the fundamental expression elements. This leads to the construction of a self-generated and hierarchical tree-structured category of GA representation.

Addressing common GIS data models, multidimensional geometric data can be decomposed into hierarchical expressions of primitive elements using outer products. Network data can be decomposed into nodes, edges, and paths expressed hierarchically through outer products, while field data can be broken down into tensor product hierarchies of varying dimensions. Due to their shared hierarchical structure, these three types of data can be categorized through the mapping functor from GA blades, which paves the way for creating a multidimensional data category graph for cross-domain expression.

Abstraction of computation categories

The construction of computation categories requires unifying diverse GIS operations and analytical workflows. Firstly, we extend the representation categories into parameter categories, including data parameters, semantic parameters, and conversion parameters. Secondly, the algebraic operators from GA can also be easily converted into operator categories, including transform operators, dimension operators, and relational operators. By combining the parameter categories and the operator categories, a unified computational template, denoted as *f*(*n*, *parameters*, *operators*), can be constructed. This template can then be further extended into template categories designed for divergence, sequential, and aggregate computation workflows. This approach realizes the operationalization and templating of computations, facilitating the extension of more complex computational processes based on these template categories.

TYPICAL COMPOSITING PARADIGMS IN CGGC

Following the establishment of representation and computation categories, category-based relation composition and mapping provide a versatile and scalable paradigm for domain applications. We will give three typical compositing paradigms: data aggregating, function extending, and task co-designing (Figure 1).

The data aggregating paradigm emphasizes utilizing a data type graph based on representation categories, which can connect diverse data source and apply natural transformations from different domains to realize dynamic updates and aggregation. The function extending paradigm is primarily realized by constructing a functor between the data category and the operator category. For example, based on the original categories of static data and static operators, we can build a functor to map static data to dynamic data and static operators to dynamic operators, thereby extending the analysis capabilities from static to dynamic scenarios. The task co-designing paradigm primarily addresses cross-domain collaborative applications. By decomposing workflows into syntax diagrams and semantic components, it separates formal construction rules from domain semantics in complex tasks. This separation facilitates clearer theoretical analysis and enables reasoning and manipulation at different levels, thereby allowing for the development of heterogeneous systems for large-scale tasks.

SUMMARY

The reductionist perspective GeoComputation played a crucial role in the initial explosive growth phase of GIS. However, as GIS evolved into stages of interdisciplinary integration and consolidation, its capability in systematic handling and structured information generation remained relatively weak. This inadequacy restricts its ability to support the unified management of cross-domain information in the era of big data and artificial intelligence. In this paper, CGGC, which builds upon the previously constructed GA data model to establish a formalized expression for GIS, is developed. It extracts the common features within each structure, and the inter-structure relationships of GIS data and function to form the data categories and operator categories. The proposed three compositional paradigms demonstrate the scalability of CGGC across diverse domains and its aptitude for cross-domain collaboration.

This research effectively advances the mathematical foundation of GeoComputation from set theory to category theory and shifts the research paradigm from a bottom-up reductionism approach to a top-down holism one. As unseen domains learning has become a hot research topic in artificial intelligence, in-depth research on CGGC can further enhance the application potential of geospatial intelligence and even quantum geographic computing, thereby comprehensively improving the analysis and application capabilities of GIS.

REFERENCES

1. Lü G, Yue S, Yu Z, et al. (2023). Ubiquitous geographic information for building digital twins of geographic environments. The Innovation Geoscience, **1**(2): 100023. DOI: 10.59717/j.xinn-geo.2023.100023.
2. Sui D. (2014). Ubiquitous Computing, Spatial Big Data and Open GeoComputation. In: Abrahart R J, See L M, eds. GeoComputation. CRC Press, Boca Raton, 375–390.
3. Frank, A. U. (2006). Practical Geometry: The Mathematics for Geographic Information Systems (unpublished), https://www.researchgate.net/publication/265272329_PRACTICAL_GEOMETRY_THE_MATHEMATICS_FOR_GEOGRAPHIC_INFORMATION_SYSTEMS.
4. Kelly G. M. (1982). Basic Concepts of Enriched Category Theory. Cambridge University Press, 1–136.
5. Yuan L, Yu Z, & Luo W. (2019). Towards the next-generation GIS: a geometric algebra approach. Annals of GIS, **25**(3): 195–206. DOI: 10.1080/19475683.2019.1612945.

FUNDING AND ACKNOWLEDGMENTS

The work is supported by the National Natural Science Foundation of China (42130103). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.