A KNOWLEDGE GRAPH MODEL FOR PERFORMANCE-BASED GENERATIVE DESIGN AND ITS APPLICATIONS IN ACCELERATED DESIGN

ZHAOJI WU[1], ZHE WANG[2], JACK C. P. CHENG[3] and HELEN H. L. KWOK[4]
[1,2,3] Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China.
[2] HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen, China.
[4] Institute for the Environment, The Hong Kong University of Science and Technology, Hong Kong SAR, China.
[1] zwubz@connect.ust.hk, 0000-0002-6739-7909
[2] cezhewang@ust.hk, 0000-0002-2231-1606
[3] cejcheng@ust.hk, 0000-0002-1722-2617
[4] hlkwokab@connect.ust.hk, 0000-0002-7179-9281

Abstract. Data integration and information enrichment pose significant challenges to the advancement of Performance-based Generative Design (PGD). One potential solution to these challenges is the utilization of Knowledge Graph (KG). However, the implementation of KG in PGD, particularly in leveraging expert knowledge to accelerate the process, remains an area that has not been thoroughly explored. In this research, we propose a PGD-KG schema to capture and represent the topological relationships and functionalities within PGD. We also introduce a method for automatically generating PGD-KG models from parametric design models enriched with semantic information. Additionally, we develop reasoning algorithms based on expert knowledge of sustainable design to facilitate automated performance evaluation. The effectiveness of the PGD-KG approach was demonstrated through its implementation in a design project, where the reasoning algorithms proved capable of significantly reducing the solution space in PGD by 88.50%, while still ensuring the inclusion of an adequate number of Pareto optimal solutions. This research contributes to the design acceleration by integration of expert knowledge, particularly sustainable design strategies, into PGD.

Keywords. Performance-based generative design, Knowledge graph, Reasoning algorithm, Building performance evaluation, Sustainable building design.

1. Introduction

The demand for green buildings has gained significant momentum, driven by the urgent need to alleviate the environmental impact of the built environment. As shown

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 395-404.

by The MacLeamy Curve (Davis, 2011), the earlier stage in a project, the less costly it is to make a change. Hence, it is important to incorporate green performance during the early design stage of buildings. Early design is a complex task which involves multiple factors, e.g., building programs, constraints, parameters, and objectives. Performance-based generative design (PGD) (or performative computational architecture (Ekici et al., 2019)) is a powerful tool to mitigate the complexity of early design with the assistance of computational methods and artificial intelligence. In the current practices of PGD, architects create and manipulate a geometric model in parametric design tools (e.g., Rhino and Grasshopper). However, to conduct building performance evaluation, more information (e.g., materials and constructions) is needed beyond geometry (Negendahl, 2015). Although some tools (e.g., Ladybug Tools) have been developed to facilitate the information enrichment in the process of PGD, architects still need to collect and input the information manually or semi-automatically, which requires expert knowledge in building evaluation. Consequently, data integration and information enrichment are two important issues that hinder the acceleration process of PGD.

A knowledge graph (KG) represents knowledge in a structured and interconnected manner, capturing the relationships between entities and properties among multiple domains (Ji et al., 2022). By providing a unified schema and semantic framework, KG enables seamless integration of diverse data sources (Pauwels et al., 2017). In addition, the graph-based representation of knowledge in KG enables automated reasoning, which can uncover implicit relationships and enrich the semantic information of the KG model (Pauwels et al., 2017). With these advantages, KG could provide a solution to the issues of data integration and information enrichment for PGD. In addition, several studies (Machairas et al., 2014; Su & Yan, 2015) have stated that expert knowledge can help accelerate PGD. For example, design knowledge from experts or architects could help minimize the size of the solution space (Machairas et al., 2014) and customize the optimization process (Su & Yan, 2015). However, how to implement KG to PGD and accelerate PGD with expert knowledge has not been fully studied yet.

The purpose of this research is to develop a KG model for PGD and apply it to the process of PGD. Our specific goals are: 1) to design the KG schema for PGD (PGD-KG) that can integrate data from multiple domains, 2) to propose the PGD-KG generation method from parametric design models with enriched semantics, 3) to integrate expert knowledge to accelerate PGD by developing PGD-KG-based reasoning algorithms for automatic performance evaluation. Finally, one detached house design project is selected as the illustrative example of the implementation of the PGD-KG, and performance of the PGD-KG in accelerated design is evaluated.

2. Methodology

2.1. PGD-KG SCHEMA

Previously, we proposed an ontology model for building energy modelling (Wu et al., 2023), which integrates four domains (weather, building, internal heat gain and Heating, Ventilation, and Air Conditioning (HVAC) system) based on the existing models, Brick Schema (Balaji et al., 2018) and Building Topology Ontology (BOT)

(Rasmussen et al., 2020). The KG model schema for PGD is developed by leveraging the previous ontology model (Wu et al., 2023), with necessary extensions implemented to comprehensively capture and represent the topological relations and functionalities in PGD. Figure 1 shows the overview of the PGD-KG schema that covers five key

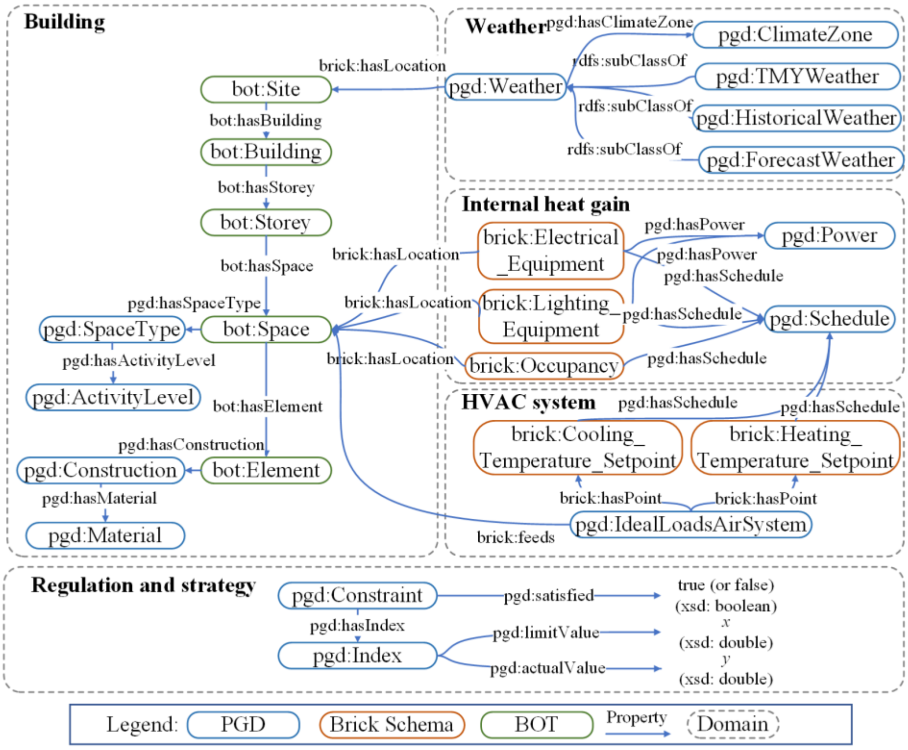


Figure 1. PGD-KG schema

domains in PGD, including building, weather, internal heat gain, HVAC system, and regulation and strategy. The components beyond the existing ontology models (Brick Schema and BOT) are assigned a prefix pgd.

Building is the important domain in the PGD-KG schema. The class bot:Site is assigned with the property pgd:has Area_m2 to support the calculation of site-area-related indices, e.g., plot ratio. The class bot:Building stores the information on building types and building directions. The class bot:Space describes spatial zones, and stores the geometric information (ceiling heights, floor areas, volumes, etc.). The topology property pgd:hasAirFlow is assigned to the class bot:Space to describe the topology relation of airflow among spaces. Each bot:Space is connected with the class pgd:SpaceType that indicates the space's function and the activity level of occupants in the space. The class bot:Element describes surfaces including building surfaces (walls, floors, roofs, ceilings, air walls, etc.), fenestration surfaces (windows, doors, etc.), and shading surfaces. Topology relations in a building surface contain the information on surface types, geometries (surface areas, vertex coordinates, normal vectors, etc.), boundary conditions (whether the surface is exposed to the outdoor environment), constructions and materials (thermal properties, optical properties, etc.). Air walls are commonly adopted to describe large openings between two spaces in modelling software, e.g., EnergyPlus. In the PGD-KG schema, air walls are regarded as a kind of building surfaces with the construction AirBoundary, which is aligned with the setting in EnergyPlus. The topology relations in a fenestration surface and a shading surface are similar to the ones in a building surface but the material properties are different. Minor revisions of the components in the domains of weather, internal heat gain and HVAC system are made compared with the previous ontology model (Wu et al., 2023) and details of these components can be referred to Section 3.1 in Ref. (Wu et al., 2023). The graphs and files of the PGD-KG model can be found on https://github.com/GeorgeZWu/PGD_KG_Schema.

2.2. PGD-KG MODEL GENERATION

Figure 2 depicts the workflow of the PGD-KG model generation. Firstly, spaces and surfaces in parametric design models are translated into individuals under the classes bot:Space and bot:Element in the PGD-KG model. Secondly, properties of the spaces and surfaces, including volumes and floor areas of spaces, types, areas, vertex coordinates and normal vectors of surfaces, etc., are identified or calculated. Corresponding KG properties (pgd:hasVolume_m3, pgd:hasFloorArea_m2, pgd:hasSurfaceType, pgd:hasArea_m2, pgd:hasVertex1X_m, pgd:hasNormalVectorX_m, etc.) are generated. Thirdly, semantic information including weather, space type, etc. is translated into individuals under the classes pgd:Weather, pgd:SpaceType, etc. of the PGD-KG model. The semantic information is derived from default templates which could be reused in various projects, or the semantic information is manually input by designers for specific cases. Finally, knowledge inference is conducted based on the generated PGD-KG model, and boundary conditions and topology relations of airflow are supplemented to the PGD-KG model.

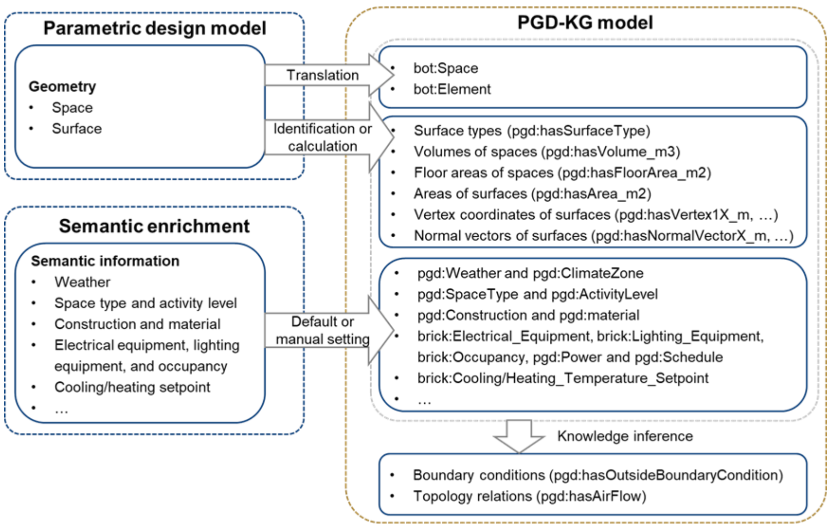


Figure 2. PGD-KG model generation from parametric design models and semantic enrichment.

2.3. REASONING FOR PERFROMANCE EVALUATION

Figure 3 elaborates the extraction workflow of design strategies based on expert knowledge from text-based literature to machine-readable reasoning algorithms. Firstly, design strategies are summarized from the literature on sustainable building design. In this research, three design strategies (summarized in Table 1) are extracted from literature (Brown & DeKay, 2013; Heywood, 2019; Kwok & Grondzik, 2018). Secondly, the design strategies are translated into machine-readable algorithms that support cross-domain reasoning of the PGD-KG models. Table 2 shows one example of the reasoning algorithm in SPARQL for the number of nodes in the overall airflow path, the indicator of Strategy 2.

Following the execution of the algorithms, the metrics for the three indicators can be derived. The sustainable performance of design cases can be evaluated based on the rankings of these indicators. To determine the number of selected cases, a hyperparameter called the threshold is introduced. For example, if the threshold is 40%, the cases with indicators ranking within the top 40% for all strategies will be chosen.

Table 1. Design strategies derived from literature.

	Interpretation	Indicator	Performance
1	The area of exterior envelopes should be as less as possible.	Shape factor	Cooling and heating
2	The ventilation path should be sufficiently short, and the resistance should be sufficiently low.	Number of nodes in overall airflow path	Cooling
3	The light-transmitting windows should be mainly directed to the south or north, and the light-transmitting windows on the south and north side should be sufficiently large.	Projection area of light-transmitting windows on the south and north side	Daylighting and cooling

Table 2. Reasoning for the number of nodes in overall airflow path.

Algorithm 1 (in SPARQL): Reasoning for the number of nodes in overall airflow path.

```
SELECT (COUNT(DISTINCT ?intermediateNode) AS ?nodeCount)
WHERE {
    ?window1 rdf:type bot:Element .
    ?window1 pgd:hasSurfaceType "Window" .
    ?window1 pgd:hasOutsideBoundaryCondition "Outdoors" .
    ?window2 rdf:type bot:Element .
    ?window2 pgd:hasSurfaceType "Window" .
    ?window2 pgd:hasOutsideBoundaryCondition "Outdoors" .
    ?window1 pgd:hasAirFlow* ?intermediateNode .
    ?intermediateNode pgd:hasAirFlow* ?window2 .
    }
```

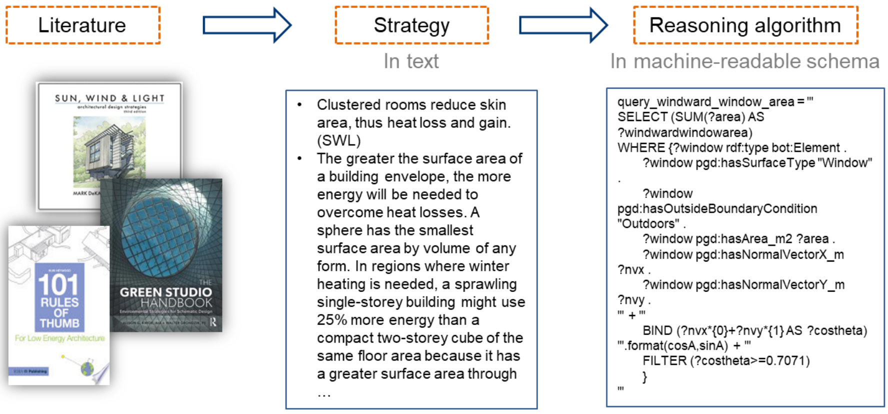


Figure 3. Extraction of design strategies from text-based literature to machine-readable reasoning algorithm.

3. Illustrative Example

3.1. INTRODUCTION OF THE DESIGN PROJECT

A 2-storey detached house using prefabricated modular techniques in a pilot design study was selected as the illustrative example of the implementation of the PGD-KG. Figure 4 illustrates the modelling rule of the prefabricated modular detached house. The sizes and locations of the staircase and living rooms were settled according to the design constants (Table 3), while the positions and orientations of other prefabricated modular units varied among five slots (A, B, C, D and E). There were two directions (along the x-axis and y-axis of the building coordinate) of the prefabricated modular units in slot A, C and D, while the prefabricated modular units in slot B and E were only directed along the y-axis. Five kinds of module combination were allocated in the five slots according to the design variables (Table 3). The goal of the design project was to generate optimal layouts of the module combinations that minimize the cooling energy consumption and maximize the daylighting performance of the detached house.

The implementation was programmed by Python. Firstly, the parametric design models of the detached house were developed using Python library rhino3dm, which offers geometric manipulation functions that can create and modify parametric Rhino 3D models in Python. Secondly, the PGD-KG models were generated based on the parametric design model and enriched semantics. Thirdly, reasoning for performance evaluation based on the three design strategies was conducted. The Python library rdflib was adopted to construct the PGD-KG models and conduct reasoning for the performance evaluation. We studied the sizes of the solution spaces selected by the reasoning algorithms for performance evaluation under different thresholds (40%, 30%, 20% and 10%) with three metrics: the ratio of selected solutions (the number of solutions selected by the reasoning algorithms divided by the total number of cases), the ratio of selected Pareto optimal solutions (the number of selected Pareto optimal solutions divided by the number of Pareto optimal solutions in the ground truth) and computational time reduction (compared with simulations of all the cases). In addition, all the cases (7680 in total) were simulated by EnergyPlus and Radiance to obtain their performance metrics Annual Cooling Load (ACL) and Useful Daylight Illuminance (UDI) (Nabil & Mardaljevic, 2006) as the ground truth of the performance.

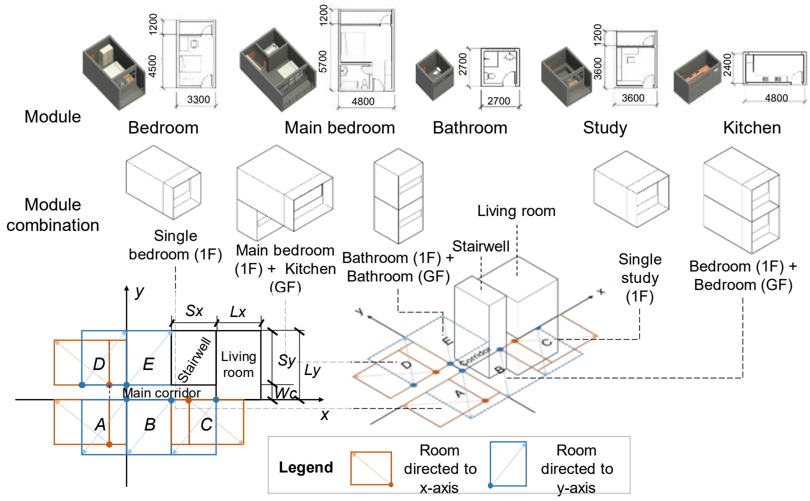


Figure 4. The modelling rule of the prefabricated modular detached house.

Table 3. Design parameters of the prefabricated modular detached house.

Parameter		Description	Value
Constant	Wc	Corridor width	1.2 m
	H	Floor height	3.3 m
	Lx	Living room width	3.0 m
	Ly	Living room depth	4.8 m
	Sx	Stairwell width	1.8 m
	Sy	Stairwell depth	3.6 m
Decision variable	Sa	Combination in slot A	Range: [0, 4], step: 1
	Sb	Combination in slot B	0: Single bedroom (1F);
	Sc	Combination in slot C	1: Main bedroom (1F) + Kitchen (GF);
	Sd	Combination in slot D	2: Bathroom (1F) + Bathroom (GF);
	Se	Combination in slot E	3: Single study (1F);
			4: Bedroom (1F) + Bedroom (GF).
	Da	Direction of the combination in slot A	Range: [0, 1], step: 1
	Dc	Direction of the combination in slot C	0: along y-axis;
	Dd	Direction of the combination in slot D	1: along x-axis.
	O	Orientation	Range: [0,360°), step: 45°

3.2. RESULTS AND DISCUSSION

The PGD-KG models of all the design cases could be generated from their parametric design models with enriched semantics. Fig. 13 shows the classes, properties, individuals, and metrics of the PGD-KG model of one example (Case 6725) that is visualized in Protégé.

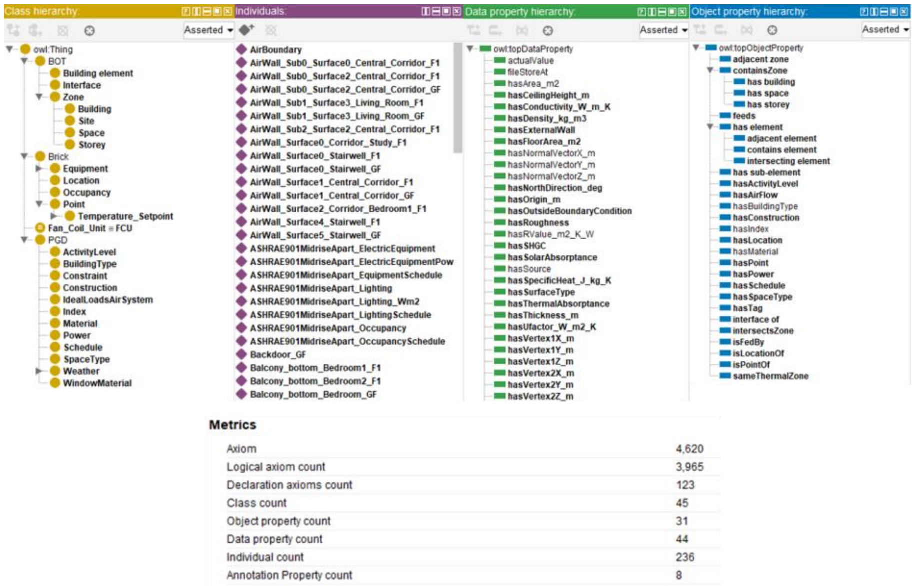
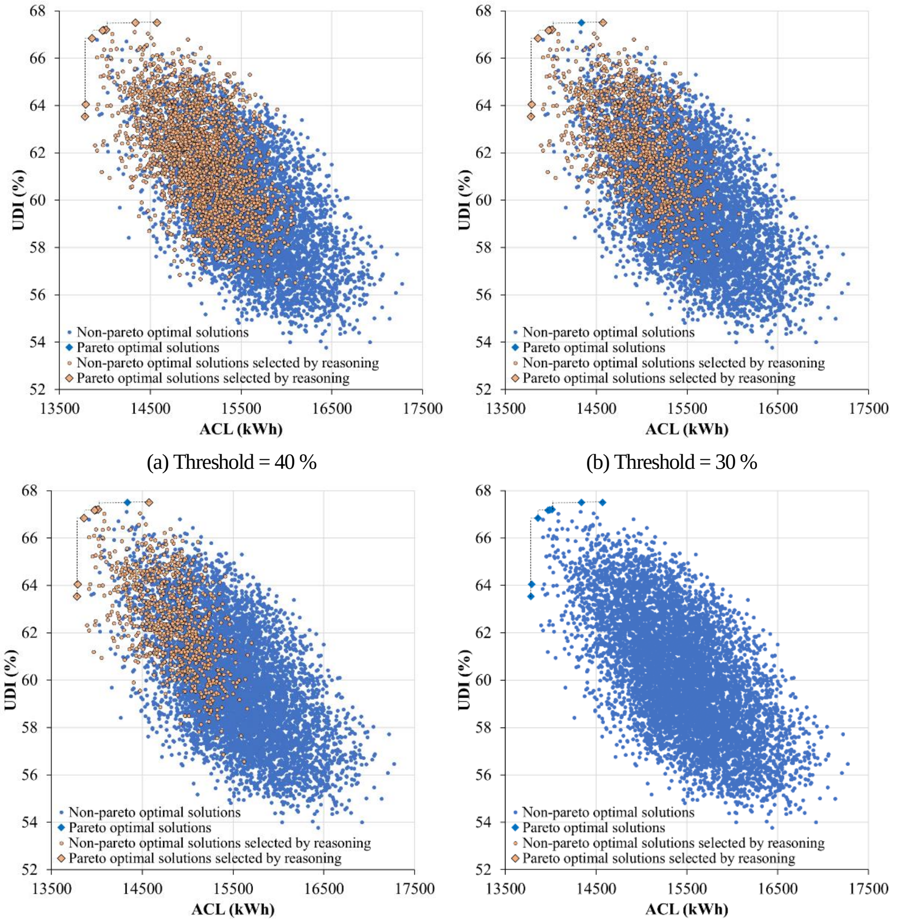


Figure 5. Classes, properties, individuals, and metrics of the PGD-KG model of Case 6725, visualized in Protégé.

Table 4. Performance of the PGD-KG reasoning under different thresholds.

Threshold (%)	Ratio of selected solutions (%)	Ratio of selected Pareto optimal solutions (Selected / ground truth)	Computational time reduction (%)
40	31.04	8/8	64.73
30	16.33	7/8	75.25
20	11.50	7/8	75.21
10	0	0/8	0

Table 4 shows the performance of the PGD-KG reasoning and Figure 5 shows the selected solution spaces by the reasoning algorithms for performance evaluation under different thresholds (40%, 30%, 20% and 10%). As the threshold was more stringent (i.e., smaller in magnitude), the ratio of the selected solutions became smaller and the selected solution space contracted, orienting towards lower ACL and higher UDI values. When the threshold was equal to 40%, all the Pareto optimal solutions were included in the selected solution space. When the threshold was 30% and 20%, one Pareto optimal solution was not included. When the threshold was 10%, no solution was selected, showing that no case satisfies all the strategies if the threshold was too stringent. In the illustrative example, it took averagely 11.0 s to conduct performance evaluation reasoning while it needed averagely 135 s to conduct performance simulation for each case. The reasoning algorithms can effectively narrow down the solution space in PGD while ensuring the inclusion of an adequate number of Pareto optimal solutions (≥ 87.5% in the illustrative example). The narrower solution space leads to less cases to be simulated and analysed, and hence time for simulations in PGD is reduced.



(a) Threshold = 40 %

(b) Threshold = 30 %

(c) Threshold = 20 %

(d) Threshold = 10 %

Figure 6. Selected solution spaces under different thresholds.

4. Conclusions

In this research, we propose a PGD-KG schema that effectively captures and represents the topological relationships and functionalities within Performance-based Generative Design (PGD). Furthermore, we introduce a generation method that automatically generates PGD-KG models enriched with instances from parametric design models, incorporating semantic information. Additionally, we develop PGD-KG-based reasoning algorithms for automated performance evaluation. To evaluate the effectiveness of the PGD-KG approach, we implemented it in a design project and observed that the reasoning algorithms successfully narrowed down the solution space in PGD, while ensuring the inclusion of a satisfactory number of Pareto optimal solutions. This research realizes the unsolved ideas proposed by the previous studies including minimizing solution spaces (Machairas et al., 2014) and customizing optimization processes (Su & Yan, 2015) by the automated integration of expert knowledge into PGD. However, only one design case is illustrated in this research. In the future, more design projects should be studied to further evaluate the efficacy and reliability of the proposed method.

References

Balaji, B., Bhattacharya, A., Fierro, G., Gao, J., Gluck, J., Hong, D., Johansen, A., Koh, J., Ploennigs, J., Agarwal, Y., Bergés, M., Culler, D., Gupta, R. K., Kjærgaard, M. B., Srivastava, M., & Whitehouse, K. (2018). Brick: Metadata schema for portable smart building applications. Applied Energy, 226, 1273–1292. https://doi.org/10.1016/j.apenergy.2018.02.091

Brown, G., & DeKay, M. (2013). Sun, wind, and light (The 2nd edition). Wiley.

Davis, D. (2011). The MacLeamy curve. The MacLeamy Curve. https://www.danieldavis.com/macleamy/

Ekici, B., Cubukcuoglu, C., Turrin, M., & Sariyildiz, I. S. (2019). Performative computational architecture using swarm and evolutionary optimisation: A review. Building and Environment, 147, 356–371. https://doi.org/10.1016/j.buildenv.2018.10.023

Heywood, H. (2019). 101 rules of thumb for low energy architecture. Riba publishing.

Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P. S. (2022). A Survey on Knowledge Graphs: Representation, Acquisition, and Applications. IEEE Transactions on Neural Networks and Learning Systems, 33(2), 494–514. https://doi.org/10.1109/TNNLS.2021.3070843

Kwok, A. G., & Grondzik, W. (2018). The green studio handbook: Environmental strategies for schematic design. Routledge.

Machairas, V., Tsangrassoulis, A., & Axarli, K. (2014). Algorithms for optimization of building design: A review. Renewable and Sustainable Energy Reviews, 31, 101–112. https://doi.org/10.1016/j.rser.2013.11.036

Nabil, A., & Mardaljevic, J. (2006). Useful daylight illuminances: A replacement for daylight factors. Energy and Buildings, 38(7), 905–913.

Negendahl, K. (2015). Building performance simulation in the early design stage: An introduction to integrated dynamic models. Automation in Construction, 54, 39–53. https://doi.org/10.1016/j.autcon.2015.03.002

Pauwels, P., Zhang, S., & Lee, Y.-C. (2017). Semantic web technologies in AEC industry: A literature overview. Automation in Construction, 73, 145–165. https://doi.org/10.1016/j.autcon.2016.10.003

Rasmussen, M. H., Lefrançois, M., Schneider, G. F., & Pauwels, P. (2020). BOT: The building topology ontology of the W3C linked building data group. Semantic Web, 12(1), 143–161. https://doi.org/10.3233/SW-200385

Su, Z., & Yan, W. (2015). A fast genetic algorithm for solving architectural design optimization problems. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 29(4), 457–469. https://doi.org/10.1017/S089006041500044X

Wu, Z., Cheng, J. C. P., Wang, Z., & Kwok, H. H. L. (2023). An ontology-based framework for automatic building energy modeling with thermal zoning. Energy and Buildings, 296, 113267. https://doi.org/10.1016/j.enbuild.2023.113267