**Abstract**

This research aims to enhance a combinatorial process for the generation of spatial assemblages with adaptive control over climate-related factors, such as daylight access, by means of Reinforcement Learning (RL).

In combinatorial design, a finite set of parts and rules for their coupling and iterative aggregation, generates larger assemblies whose properties, performances, and functions, at different system scales, differ from those of the constituent parts; the whole is composed of multitudes, engendered by the mutually occurring interactions. A key role in establishing the potential for the emergence of holistic qualities lies in the design of both parts and rules as well as in the policy that regulates rule selection in relation to the design goals.

Combinatorial design is defined by Sanchez as an inherently open process, where no kind of optimization is achievable. However, it is possible to operate within this context to create systems whose holistic properties, related to quantitative aspects of the architectural space, follow predetermined design criteria. In their combinatorial research, both Sanchez, Alexander (in “*A Pattern Language”),* and Stiny (in his study on “*Shape Grammars”)* use fixed criteria (either stochastic or heuristics-based) for the choice of aggregation; a different strategy is using a reward-oriented policy to derive the aggregation rule to apply at each iteration.

From this perspective, Makoto Sei Watanabe in his *Induction Cities* series experimented with "inductive" models: a wide range of viable spatial configurations generated via a stochastic process coupled with selective target conditions based on climatic factors such as direct sunlight access (*Sun God City).*

Building on Watanabe's work, this study aims to bind combinatorial logic with topological considerations and environmental-climatic feedback, for the context-adaptive generation of spatial units assemblages with control over daylight hours access on exposed surfaces. The approach uses RL-trained agents instead of random choice/heuristics iterative algorithms to adapt to lighting conditions in the process of selection and aggregation of parts. The agents move and place parts (voxels) within a voxelized space, aiming to ensure topological consistency and a target daylight hours access on exposed surfaces in the final assembly.

The research introduces condition-based adaptivity into a combinatorial process by means of RL training, moving beyond both random choice and predetermined heuristics sets; although both can relate to boundary conditions, they are respectively non-controllable and tied to a specific environmental scenario. Through the agent’s trained policy, the system learns a state-action-reward relationship in a process of continuous feedback between space, environment and climate data that applies to any environmental configuration that can be coded in the system’s terms.

The study is implemented coupling state-of-the-art Python RL libraries (Stable Baselines 3, Gymnasium) and the Rhino+Grasshopper environment for modelling, daylight factor calculation, and visualization, building a custom infrastructure for bidirectional data communication between computing environments during the training and inference phases.