Artificial Intuitions of Generative Design: An Approach Based on Reinforcement Learning

Dasong Wang and Roland Snooks^(✉)

Royal Melbourne Institute of Technology (RMIT), Building 100, Melbourne 3000, Australia
{dasong.wang, roland.snooks}@rmit.edu.au

Abstract. This paper proposes a Reinforcement Learning (RL) based design approach that augments existing algorithmic generative processes through the emergence of a form of artificial design intuition. The research presented in the paper is embedded within a highly speculative research project, Artificial Agency, exploring the operation of Machine Learning (ML) in generative design and digital fabrication. After describing the inherent limitations of contemporary generative design processes, the paper compares the three fundamental types of machine learning frameworks in terms of their characteristics and potential impact on generative design. A theoretical framework is defined to demonstrate the methodology of integrating RL with existing generative design procedures, which is further explained with a Random Walk based experimental design example. The paper includes detailed RL definitions as well as critical reflections on its impact and the effects of its implementation. The proposed artificial intuition within this generative approach is currently being further developed through a series of ongoing and proposed research trajectories noted in the conclusion. The ambition of this research is to deepen the integration of intention with machine learning in generative design.

Keywords: Generative design · Machine learning · Reinforcement Learning · Intelligent formation

1 Introduction

Architectural design has been fundamentally impacted over the past three decades by the integration of emerging technologies and processual theory which have contributed to the proliferation of generative design methodologies [1]. Among these, the rapid maturity of artificial intelligence techniques, the massive increase in computational power and further development of complexity theory provides a new perspective to critically reflect on future directions of generative architectural design [2]. Framed by the limitation of contemporary generative methodologies, this paper proposes a Reinforcement Learning based approach to integrate Machine Learning with current computational design processes. This will be demonstrated through a simple design experiment based on a random walk algorithm. The broader ambition of this research is to leverage the generative potential of contemporary processes while integrating intuition through machine learning.

1.1 Contemporary Algorithmic Generative System

Contemporary generative design approaches can be categorized into two broad types, roughly summarized as parametric-based and behavioral-based. Parametric design processes operate through the manipulation of parameters that have an established linear relationship to a set of known geometric procedures. While, in behavioral-based systems, the control operates through encoding design intentions into a series of local behaviors to form a bottom-up, self-organizing process [3].

While both approaches are capable of compelling and sophisticated design outcomes, a series of limitations, or bottlenecks, still exist as obstacles towards its further development. Initially, the parametric-based approach relies on the linear relationship between parameters and the system/geometry, which leads to models that are limited to their predefined conditions. While, the behavioral-based system privileges micro interactions over macro awareness [4], establishing a global ignorance that limits the integration of overall design intentions. Furthermore, the integration of real-time materialization and structural performance [5] within non-linear generative design processes remains problematic due to the inherent volatility of these methodologies.

1.2 Artificial Intuitions

The paper speculates on a generative process driven by machine learning, which is capable of gradually developing typical and specific artificial “intuition” towards a series of design intentions. In natural processes of evolution, intuitions emerge from intelligent creature’s inheritance of long-time accumulation of knowledge from generation to generation. The approach posited in this paper is intended to form a higher level of (machine) intelligence within generative design by undertaking a self-training, learning, and incrementally evolving process.

The research presented in the paper is part of an ongoing research project, Artificial Agency, with aims to explore the operations of machine learning with generative design and autonomous fabrication process, undertaken at the RMIT Architecture, Snooks Research lab.

2 Background

2.1 Machine Learning with Generative Design

The proposed intuitive generative approach is inspired by, and based on, the development of machine learning techniques. Contemporary machine learning consists of three fundamental types of frameworks: supervised learning, unsupervised learning and reinforcement learning [6] (Fig. 1).

Supervised Learning (SL) is essentially an algorithm that trains a predictive model with a labelled dataset (known outcomes). In recent years, enormous progress has been achieved with the rapid development of SL across a wide range of fields: data-prediction, image-synthesis, language-processing, etc. [7]. However, the impact of SL on three-directional generative design is still to be explored. Firstly, SL relies on massive labelled dataset, which is considered as a highly inefficient [8] and unrealistic process. While

Fig. 1. Diagram of three typical types of machine learning: Supervised Learning (SL), Unsupervised Learning (UL), Reinforcement Learning (RL).

the labelling operation could be undertaken algorithmically, the feedback from the two sides of the ANN (Artificial Neural Network) is a linear procedure regardless of the data parsed during the generating process, which is opposed to the ambitions of existing generative design. Additionally, three-dimensional geometry representations are problematic with SL, and in particular with 3D GAN (Generative Adversarial Network) algorithmic frameworks [9], due to the substantial computational requirements.

Comparatively, Unsupervised Learning (USL) is based on training a clustering and association model with non-labelled datasets. Generally, USL doesn't have a clear training objective, but instead it aims to uncover invisible relationships within a massive dataset. Consequently, this approach is problematic when working with generative approaches that involve specific design intention.

2.2 Reinforcement Learning

Reinforcement Learning (RL) is closely associated with the field of optimal control, in which an agent seeks an optimal policy by interacting with its environment through a feedback between observation states and quantified rewards, modeled as a Markov Decision Process [10] with following specific elements (Fig. 2).

Fig. 2. Diagram of Reinforcement Learning (RL) with main elements: agent, environment, state, reward, action.

- **Observation State (S):** State is a concrete and immediate information summary of the agent itself and its interaction with the environment.
- **Agent Action (A):** Action is a set of possible moves the agent can take to interact with the environment.
- **Reward (R):** Reward is the feedback that measures the success or failure of an agent's actions in a given observation state.

- Policy (π): Policy is the strategy that the agent employs to determine the next action based on the current state. It maps states to actions, undertaking the actions that return the highest reward.

Under the overall structure of RL, there are diverse implemented algorithms: Q-Learning (Value-Based), Policy Gradient (Policy-Based), Actor-Critics, as well as further research fields: Hierarchical RL, Multi-Agent RL, etc. Contemporary RL has achieved significant progress with its application in Gaming AI, Self-Driving Vehicles and Robotics fields since 2017 [11].

It can be seen that RL has a clear correlation with, and enormous potential impact on existing generative design processes. Firstly, RL operates in a heuristic mode with no direct human knowledge, as opposed to the labelling process of SL. This heuristic mode is conceptually similar to the objective of generative design: to create the unpredictable and previously unimagined through logical design intentions. Secondly, RL operates on a sequential decision tree rather than the simultaneous processing of massive datasets (SL), which is suitable to be implemented with the constantly evolving generative controlling process. Thirdly, there are multiple technical approaches to implementing RL within generative design in three-dimensional environment, such as Gym toolkits [12] by OpenAI and ML-Agents toolkits [13] within Unity3D platform.

3 Methodology

The framework of the proposed design approach is to integrate RL with existing generative processes, in which RL is acting as a brain to further control the algorithmic system instead of creating an entirely new procedure. The methodology is further demonstrated with a Random Walk based design experiment from the overall training setup to detail definitions.

3.1 Intuitive Random Walk Formation

Random Walk (RW) is a long-standing algorithmic model inspired by a natural stochastic process [14], with applications in numerous scientific fields. As shown in Fig. 3, the goal of this example is to train a RW with a series of basic architectural intuitions, initially inspired by Le Corbusier's Domino System [15] and further developed with more abstract and critical design intentions of spatial and structural logic. With the implementation of RL, it is expected that the opposing characteristics of Random Walk's stochastic operation and the Domino System's formality can be integrated with a synthetic design process.

3.2 RL Actions Definition

Within the training framework, RL actions can be based on the underlying generative system or customized methods to further control the generating process, depending on the characteristics of the system and training task (Fig. 4).

Fig. 3. Diagram demonstrating the training intention of the Domino system as a type of intuition within the RW generative formation process.

Fig. 4. Diagram of the definition of RL actions within the RW generative process.

The action definition of the RW experiment is that an agent takes random decisions to move towards six directions within a limited three-dimensional voxel grid, from which the walking trail is recorded as a generated form. In this case, the RL action is considered as a discrete [10] action, with a vector¹ size of seven.

3.3 RL Observations Definition

The definition of observation states describes the current condition within the generating process, which normally consists of two types of information: the overall matrix data type representation of the form and a series of significant reward-oriented values (Fig. 5).

Fig. 5. Diagram of the definition of RL observations within the RW generative formation process.

In the RW example, the form is converted to a three-dimensional representation: voxel-based matrix of integers (1 or 0), representing a Boolean describing whether the

¹ Vector: refers to a one-dimensional array (being distinguished from vector definition of Computer Graphics) .

voxel is occupied or not. Additional reward-oriented information is also included in the states, such as the current position of an agent, and its real-time reward evaluation figures.

3.4 RL Reward Definition

As the most critical part of the RL training process, the reward definition is normally a quantitative evaluation structure based on design intention. In this case, the initial reward definition is simply identifying some reward locations (representing domino floors) in the voxel grid and encouraging the walker to seek and connect the floors. With further development, a more comprehensive structure is setup with more detailed design intentions, showed in Fig. 6.

Fig. 6. Diagram of the definition of RL rewards within the RW generative formation process.

- **Tower Type Reward (R1):** Agent is encouraged to generate a tower-like form. The reward calculation is based on the height of the form.
- **Structural Logic Reward (R2):** A pyramid-like structural logic is implemented such that a reward at the bottom part should be larger than the stacked part above.
- **Spatial Connectivity Reward (R3):** Horizontally, if one generated voxel is connected to its four neighbouring voxels, the agent will receive a positive reward of spatial connectivity.
- **Spatial Creation Reward (R4):** The greater the void space generated in between two voxels in the vertical direction, the greater the positive reward the agent receives.
- **Site Response (R5):** Some existing voxels are setup in the grid to represent site context. When the agent collides with these voxels, a negative value will be added to the reward calculation as a form of punishment.

Fig. 7. Selected training outcome samples from episode 0 to 9.4 k.

4 Discussions

4.1 Training Process and Outcomes

The Random Walk design experiments operates with a customized Deep Q-Learning Algorithm in Python and Tensorflow Environment. Totally, the training process undertaken 10,000 episodes, calculating on a local computer with a time consumption of about three hours. In order to assess training outcomes, the generated form is recorded every 100 episodes, shown in Fig. 7. Overall, the training result is remarkably successful. The intense and squeezed form that resulted from the initial episodes (0.0 to 1.9 k) are significantly improved and evolved in the latter iterations (7.7 k to 9.4 k) in terms of the predefined reward.

The characteristics of forms generated through this process evolved unexpectedly over time, creating a clear sequence of design intentions. From episodes 0.0 to 1.9 k, the Random Walker System doesn't generate any effective intuitions. However, from 2.0 k to 4.0 k, it starts to understand the predefined intention as tower type (R1). Episodes 4.0 k to 5.9 k demonstrate how the form balances structural performance (R2) and tower type (R1) reward. Throughout the RL process the response to design intention of space connectivity (R3) and creation (R4) are slowly improved, becoming more obvious in the final episodes. The site response (R5) was not implemented in this case due to the resolution and complexity of the particular intention. Despite that all the rewards operate simultaneously and were defined prior to the training being launched, the process generates a clear, multi-stages characteristic that achieves one significant reward prior to addressing the others through a gradual process of improvement (Fig. 8).

Fig. 8. Visualization prospective of a preferred training result. (The generated voxels have been further developed into structures and platforms with a set of algorithmic stochastic operations.)

There are some obvious limitations in the posited RW design experiment that result from using a single walker to generate form within a low-resolution grid. However, as an early and speculative case to explore and demonstrate the potential approach of applying RL within generative design processes, it still shows concrete effects and significant flexibility to be deeply integrated with other existing generative design processes.

4.2 Further Research

A number of ongoing research trajectories have emerged from the posited application of RL in algorithmic generative design and digital fabrication, which are summarized as follows:

- **Complex Generative System Training with RL:** The research is focused on integrating RL with a complex self-organized generative system in response to non-programmable design intentions, such as the architectural typology logic.

- **Multi-Agent Global Awareness Training with RL:** The research aims to generate global intuitions for multi-agent systems that combines with their logic of local interactions. These global concerns include the control of form, topology and structural networks.
- **RL with Real-time Robotics:** As a collaborative direction, the research intend to apply RL with real-time robotic behavior in order to advance the concept of automated assemblies.

5 Conclusions

The proposed RL based design approach integrates heuristic design intuitions within known algorithmic generative processes, augmenting the processes to establish a greater level of sophistication and design capacity. Both a theoretical foundation and technical methodology are presented in this Intuitive Random Walk Formation case to demonstrate the concrete effects and potential flexibility of cultivating intuitions for generative systems. The subsequent reflections in the paper aim to indicate potential ways of applying this emerging tool to existing design methodologies as well as anticipating a closer correlation of designer and computational intelligence.

References

1. Frazer, J.: *An Evolutionary Architecture*. Architectural Association (Themes VII), London (1995)
2. Leach, N.: Design in THE age of artificial intelligence. *Landsc. Archit. Front.* **6**, 8–20 (2018)
3. Leach, N., Snooks, R.: *Swarm Intelligence: Architectures of Multi-agent Systems*. Tongji University Press, Shanghai (2017)
4. Snooks, R.: *Behavioral Formation: Multi-Agent Algorithmic Design Strategies*. RMIT Ph.D. thesis (2014)
5. Hensel, M., Menges, A., Weinstock, M.: *Emergent Technologies and Design: Towards a Biological Paradigm for Architecture*. Routledge, Abingdon (2013)
6. Mitchell, T.: *Machine Learning*. McGraw-Hill Science/Engineering/Math (1997)
7. Paliouras, G., Karkaletsis, V., Spyropoulos, C.: *Machine Learning and Its Applications*. Springer, Heidelberg (2001)
8. Zheng, H.: Form finding and evaluating through machine learning: the prediction of personal design preference in polyhedral structures. In: *Proceedings of the 2019 Digital FUTURES (CDRF 2019)* (2019)
9. Wu, J., Zhang, C., Xie, T., Freeman, W., Tenenbaum, J.: Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling. [arXiv:1610.07584](https://arxiv.org/abs/1610.07584) [cs.CV] (2016)
10. Sutton, R., Barto, A.: *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge (1998)
11. Henderson, P., Islamm, R., Bachman, P., Pineau, J., Precup, D., Meger, D.: Deep reinforcement learning that matters. In: *The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)* (2018)
12. Brockman, G., Cheung, V., Pettersson, L., Schneider, J., John, S., Jie, T., Zaremba, W.: OpenAI Gym. [arXiv:1606.01540v1](https://arxiv.org/abs/1606.01540v1) [cs.LG] (2016)
13. Juliani, A., Berges, V., Teng, E., Cohen, A., Harper, J., Elion, C., Goy, C., Gao, Y., Henry, H., Mattar, M., Lange, D.: Unity: A General Platform for Intelligent Agents. [arXiv:1809.02627](https://arxiv.org/abs/1809.02627) [cs.LG] (2018)

14. Kaye, B.: A Random Walk Through Fractal Dimensions. Wiley-VC, New York (1994)
15. Corbusier, L.: Vers une Architecture (Towards a new Architecture). France (1923)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.