CONTROL OF MORPHOLOGICAL VARIATION THROUGH POPULATION BASED FITNESS CRITERIA

MOHAMMED MAKKI\textsuperscript{1} and MILAD SHOWKATBAKHS\textsuperscript{2}
\textsuperscript{1,2}Architectural Association
\textsuperscript{1,2}\{mohammed.makki|showkatbakhsh\}@aaschool.ac.uk

Abstract. A primary challenge for the application of an evolutionary process as a design tool is the ability to maintain variation amongst design solutions while simultaneously increasing in fitness. The ‘golden rule’ of balancing exploration versus exploitation of solutions within the population becomes more critical when the solution set is required to present a controlled degree of phenotypic variation but ensure that convergence of the solution set continues towards increased levels of fitness. The experiments presented within this paper address the control of variation throughout the simulation by means of incorporating a population-based fitness criterion that is utilised as a fitness objective and is calculated dynamically throughout the algorithmic run in both single and multi objective design problems.

Keywords. Architecture; Computation; Evolution; Urban; Variation.

1. Introduction: The computational application of natural evolutionary processes as a problem-solving tool has been well established since the mid-20th century. However, its application within architecture and design has only gained ground in recent years, with an increasing number of academics and professionals in the field electing to utilize evolutionary computation to address problems comprised from multiple conflicting objectives with no clear optimal solution.

Recent advances in computer science and its consequent constructive influence on the architectural discourse has led to the development of multiple algorithmic processes within 3d design software capable of simulating the evolutionary process in nature within an efficient timescale (most prominently the grasshopper add-ons Galapagos and Octopus). Many of the developed processes of generating a population of candidate solutions to a design problem through an evolutionary based stochastic search process are often driven through the application of both environmental and architectural parameters, allowing for conflicting objectives to be simultaneously, independently and objectively optimized for, an approach that is essential in design problems with a final product that must address the demand of a multitude of individuals with various requirements.

However, one of the main challenges encountered through the application of an evolutionary process as a design tool is the ability for the simulation to maintain...
variation amongst design solutions in the population while simultaneously increasing in fitness. This is most commonly known as the ‘golden rule’ of balancing exploration versus exploitation over time (Luke, 2014); the difficulty of achieving this balance in the simulation is due to the tendency of either variation or optimization being favoured over the other as the simulation progresses. In such cases, the generated population of candidate solutions has either converged very early in the simulation or has continued to maintain high levels of variation to which an optimal set could not be discerned; thus, providing the user with a solution set that has not evolved efficiently to the objectives outlined in the problem at hand. As such, control over directing the degree of variation across the generation as well as among the population becomes critical for the user, this is more so within design as the morphological attributes of the resulting phenotypes are central to a successful algorithmic run.

2. Variation Through Evolutionary Computation

David Goldberg, one of the seminal figures in the field of evolutionary computation, and in specific, multi objective evolutionary algorithms, put forward the concept of integrating Pareto optimality and dominance as a selection strategy within an evolutionary algorithm, allowing for the algorithm to incrementally increase the fitness of the solutions for each fitness criteria independently yet maintain an adequate degree of variation among the population to avoid early convergence towards a local optima (Goldberg, 1989). Inspired by Goldberg’s research, many of the leading evolutionary algorithms of the 1990’s incorporated his selection strategies, most famous were the Multi-Objective Genetic Algorithm (MOGA) (Fonseca and Fleming, 1993), Niched-Pareto Genetic Algorithm (Horn et al., 1994) and the Non-Dominated Sorting Genetic Algorithm (NSGA) (Srinivas and Deb, 1994). Goldberg’s selection strategies (and incorporation thereof) advanced the field significantly; more importantly, it led to one of the field’s major advancements in the 21st century by means of introducing the concept of Elitism, a strategy primarily credited to Eckart Zitzler through his algorithm titled Strength-Pareto Evolutionary Algorithm (SPEA) (Zitzler, 1999) (the SPEA was developed into a second more robust algorithm titled SPEA-2 (Zitzler et al., 2001)). The objective of utilizing an elitism strategy (or what is sometimes called an Archive) within evolutionary algorithms was to allow non-dominated solutions to compete with individuals that lie outside of their respective generations. Zitzler’s concern was that although a non-dominated solution may have earned its non-dominated status within its own generation, it may also be non-dominated across multiple generations, however by not allowing it to ‘survive’ in order to compete with future generations, the solver may lose potentially highly fit individuals, therefore the elite were the solutions that were preserved across multiple generations and only replaced by fitter non-dominated solutions (Zitzler, 1999). Zitzler’s SPEA inspired other evolutionary algorithms to incorporate the elitism strategy, most notably Knowles and Corne’s Pareto-Archived Evolution Strategy (PAES) (Knowles and Corne, 2000) and Kalyannoy Deb’s second attempt at his NSGA algorithm titled NSGA II (Deb et al., 2000); all with the intent to better control
variation among the population while simultaneously increasing its fitness levels, thus gaining more control over the balance between exploration and exploitation of the problem’s fitness landscape.

3. Experiment Setup

3.1. CONTROLLING VARIATION AMONG THE POPULATION:

As mentioned, one of the primary challenges of evolutionary algorithms is establishing an ideal balance between the exploration and exploitation of fitness values across the population. However, achieving this balance becomes highly challenging as the problem - and the fitness landscape associated with the problem - increases in complexity. A complex fitness landscape decreases the chances of the simulation from finding a global optimum; however more importantly, it also increases dramatically the risk of the simulation generating a solution set that is restricted to a local optimum (Figure 1). As such, the experiments presented within this paper examine the possibility of utilising a population-based fitness criterion as a secondary unit of control that directs the balance between exploration and exploitation of individuals throughout the simulation.

![Figure 1. Examples of different fitness landscapes recreated from (Luke, 2014), each carrying a degree of complexity that challenges the simulation from finding a global optima.](image)

3.2. POPULATION BASED FITNESS CRITERIA

To control variation amongst the population, the differences in fitness values between all individuals within a single generation needs to be calculated. The challenge arises by the fact that this calculation must occur iteratively at the end of each generation, with the resulting value to be attributed to each solution in the generation and utilised as a fitness criterion to evolve subsequent generations. As such, a multi objective approach is needed due to the population-based fitness criterion being utilised as an additional fitness objective. Initial attempts to incorporate the population-based fitness criteria were carried out in ‘Octopus’, a grasshopper plugin that runs a multi objective evolutionary algorithm (the SPEA-2 algorithm) developed by Robert Vierlinger (Vierlinger 2013); however, as the algorithm’s loop process within ‘Octopus’ could not be interrupted, the software
‘Octopus.explicit’ - a variation of ‘Octopus’ developed by the same author - was utilized as an alternative, which allows the user to interrupt the algorithm’s workflow and make adjustments as required (Figure 2).

When calculating the population-based fitness criterion, the value necessitated two critical properties: a) The value must be derived from all the individuals within a generation, and b) it had to be a value that would be unique to each solution. Although the standard deviation value of each generation reflects the degree of variation between solutions, the value calculated is a single value, one that was not unique to each individual. As such, to bypass this issue, the deviation of each individual’s fitness value from the population average was calculated and attributed to each individual uniquely. Although this indicates the level of variance within the population, it also allows for two solutions with different fitness values that are equidistant from either side of the population average to have the same population-based fitness value, thus driving the algorithm to minimise (or maximise) variance levels by reducing (or increasing) each solution’s distance to the average in both its positive and negative. An increase in this population-based fitness criterion throughout the simulation translates to greater deviations between each individual and the generation average, meaning greater variation among solutions. In contrast, a decrease in this value conveys lower deviation of individuals to the average thus translating to lower diversity among solutions. The objective is to increase/decrease the population-based fitness criterion while simultaneously optimizing for the fitness objective used to calculate it.
4. Results:

4.1. COMPUTATIONAL SETUP

The computational setup for the design experiments presented have been developed according to the complexity of the problem being investigated. The ultimate goal of the presented experiments is the analysis and examination of the effects of incorporating a population-based fitness criterion on the morphological variation within a population. Therefore, the experiments were designed to ensure a full understanding of their results. Experiment 1 investigates the morphological variation of phenotypes within the context of an urban tissue; therefore, the selected primitive was based on the block typology of the city of Fez in Morocco, a typology that is highly integrated to its environmental and cultural context. However, within the scope of the presented experiment, the typology was modified to allow for elevated connections between multi-level spaces throughout the urban superblock (this is part of current body of research by the authors that investigates variation of the Fez superblock as an urban typology) (Figure 3).

Figure 3. Fes El Bali is distinctive in the preservation of its ancient urban configuration. The experiments presented modify the block to allow for upper-level connections.

In contrast, and in response to the complexity of the problem, the second experiment utilised a highly simplified phenotype as the base primitive (explained in the following sections), this was in response to the complex fitness landscapes that accompany multi objective problems. However, both experiments shared the same algorithmic parameters; a population size of 20 individuals per generation, a mutation rate of 50%, mutation probability of 20%, a crossover rate of 80% and an elitism size of 50%. Finally, the simulation time for both experiments was set to 50 generations (for a detailed description of the solver parameters listed above, please see (Makki et. al, 2015)).

4.2. EXPERIMENT 1 - VARIATION FOR A SINGLE OBJECTIVE

Within the first experiment, the population-based fitness criterion is applied to a single objective problem while experiment 2 (presented in the following section)
applies it to a multiple objective problem. The primitive phenotype for the first experiment was the urban superblock discussed previously. The algorithm was set up to maximise the floor areas and connections of all upper level spaces.

Both experiments 1 and 2 were carried out as a two-step process; firstly, the simulation ran without the population-based fitness criterion, in essence the algorithm was simply attempting to increase the area of upper level spaces without any restrictions; as expected, this resulted in a solution set that converged quite easily to phenotypes that had maximized upper level areas (Figure 4). Although unsurprising, this was necessary for a comparative analysis to the second step of this two-step process. Step 2 ran the exact same experiment as step 1, with the only difference being that the population-based fitness criterion was introduced, and the simulation was required to increase its value; in an attempt to reduce convergence and maintain a degree of variation among the population. The comparisons between the normal distribution curves and their respective standard deviation values as well as the generated phenotypic morphologies present very promising results (Figure 5).

The utilisation of the population-based fitness criterion as a fitness objective allowed the simulation to maintain a high degree of variation among solutions; more importantly however, the fitness levels of the generations throughout the simulation continued to incrementally increase. The results present that a population-based fitness criterion, that is derived from the values of an objective of which the algorithm is optimising for, has been able to maintain an adequate level of variation without limiting optimisation. This is examined further in the following section.
4.3. EXPERIMENT 2 - VARIATION FOR MULTIPLE OBJECTIVES

Applying the population-based fitness criterion on a problem with multiple objectives becomes significantly more challenging. The difficulty does not arise from its application, but from its analysis. Multiple objectives (and their effects on the simulation) require a highly simplistic problem in order to comprehensively examine and assess the effects of incorporating the population-based fitness criterion within the algorithm. This is due to the fact that a complex problem increases the complexity of the fitness landscape dramatically, adding multiple additional variables that may affect the balance between exploration and exploitation of solutions throughout the simulation; thus the analysis of the effects of the population-based fitness criterion becomes more difficult to discern as multiple other factors are involved in the variation and/or convergence of the population. As such, rather than relying on the primitive utilized in experiment 1 (the urban superblock); experiment 2 employed a simplified primitive derived from a 4x4 grid of blocks. The objectives defined for the simulation were the following; the algorithm will minimise the volume of the solutions while simultaneously maximise their surface envelope area - 2 criteria that are in clear conflict.

As with experiment 1, experiment 2 was carried out as a two-step process. Step 1 ran the simulation without the incorporation of the population-based fitness criterion; the result was typical to a multi-objective problem with conflicting criteria; the Pareto front distribution was concave (please note, the directions of the x and y-axes on the graph are reversed), implying that as the solution’s fitness value for one objective was optimized, this resulted in the decrease in fitness of the second objective (Figure 6). This is also visible through the normal distribution graphs for the two criteria, as they are inversely proportional to one another. However, what was unexpected (as well as unintentional) was that one criterion was favoured over the other, in this case, maximising the surface envelope
area was awarded more weight by the algorithm over minimising the total volume.

In step 2 of the experiment, the population-based fitness criterion was derived from the volume fitness objective value, therefore the algorithm setup attempted to maximise surface envelope area, minimise total volume area and minimize the variation of individuals with regards to their volume; by doing so, the ambition was to drive the experiment towards awarding greater weight in minimising the volume of solutions in an attempt to counteract the dominance displayed by the surface envelope area criterion that was exhibited in the first step of the experiment. As with Experiment 1, the results are promising; through applying the population-based fitness criterion and minimizing its value, the simulation favoured optimising the volume criterion over the surface area criterion (Figure 7). This is evident when comparing the normal distribution graphs and standard deviation factors of the populations between these two experiments, as well as the distribution of solutions in the objective space and the resulting phenotype morphologies.

![Figure 6. Evolution of simplified phenotype towards minimizing volume and maximizing surface envelope area.](image)

Although promising, significant work needs to be carried out to further develop the applied strategy and gain more control over the degree by which variation and convergence is directed throughout the simulation, however the preliminary results signify that the selected population-based fitness criterion is advantageous to controlling variation among the population within both single objective evolutionary algorithms as well as multi objective evolutionary algorithms. More importantly, there is an opportunity to investigate the benefits of activating and deactivating this criterion dynamically throughout the algorithmic run, thus serving as a regulator for when the simulation encounters high degrees of convergence or variation. This is currently being researched by the authors.
5. Conclusions:

The utilisation of evolutionary algorithms as a problem-solving strategy has been proven to be an advantageous approach for complex problems across a range of different disciplines as well as a multitude of scales. In design, the potential of evolving a population of design solutions that vary in morphological diversity is central for when the end user of the proposed design is not limited to a single individual or group of individuals, rather a population of individuals that all hold a stake in the final output. This is most prominent in design problems at an urban scale, where variation takes precedence over repetition. The generic city that is comprised from an array of the same block (or a slight variant of a block), taking into little account the geographic or environmental context, has been implemented in different parts of the world throughout the 20th century. However, with the rapidly changing climatic conditions, the exponentially growing global demographic and unprecedented migration of the population from rural areas to urban settlements, the ‘generic cities’ of the past have been unable to adapt to these changes; thus, morphological variation of blocks and superblocks within an urban tissue is paramount. As such, the degree of variation generated when utilising evolutionary algorithms is essential for allowing greater flexibility when responding to the fitness objectives driving the design experiments. However, unlimited variation serves very little purpose, but when coupled with optimization, the evolved solution set is a robust and powerful alternative to a user preference-based approach.

Control over the degree of variation amongst solutions throughout the evolved population is critical (exploration vs. exploitation). By incorporating a population-based fitness criterion that is derived dynamically from the objective values being optimised for within the algorithm, variation of solutions among the
population is directed by the levels of convergence and variation resulting from the algorithm itself. More importantly, there is an opportunity for the degree of variation to be adjusted dynamically throughout the algorithmic run, serving as a control mechanism that is suppressed in its default state, and expressed only when needed. Additionally, at its current state, solutions located on either side of the average are given the same population-based fitness value, which drives the algorithm to minimise (or maximise) the distance between the solutions on both sides of the average; however, there is an opportunity to drive the population towards increased (or decreased) average levels by minimising (or maximising) the distance between solutions on only one side of the average, thus allowing for added weight to be applied to solutions with greater (or lower) fitness values. In addition to its computational load from being minimal (there were no significant differences in computation time when compared to MOEA’s that did not incorporate the population-based fitness criterion), the practicality of utilising a dynamic population-based fitness criterion becomes highly beneficial when attempting to develop an adaptable model that can be utilised across a range of different design problems.

References