Generating Biomorphs with an Aesthetic Immune System†

Dennis L. Chao and **Stephanie Forrest**

Department of Computer Science
University of New Mexico
Albuquerque, NM 87131 USA

Abstract

We describe an interactive search algorithm inspired by the immune system. The algorithm learns what parts of the search space are *not* useful to help a user explore large parameter spaces efficiently. The algorithm is capable of finding consensus solutions for parties with different selection criteria. A simple implementation of the algorithm applied to selecting Biomorphs (Dawkins 1986) is presented.

Introduction

In this paper we illustrate how an immune system-inspired filtering algorithm described in (Chao & Forrest 2002) can be applied to the filtering and generation of art. When the immune system is exposed to a novel pathogen, it learns to recognize it in what is called the *primary response*, and it remembers the pattern so that subsequent infections can be eliminated more efficiently in the *secondary response*. Analogously, an artificial immune system can be viewed as "protecting" users from undesirable data. We propose constructing an artificial immune system that learns what kinds of data patterns are unacceptable and then prevents exposure to similar data in the future. This approach scales to multiple users by using the superset of the undesirable data of many users to filter data. The data that survive this censoring is the set that could satisfy everyone. These solutions are useful for situations in which one solution must satisfy several people at once, such as music that is broadcast to an audience or artwork that is displayed in public spaces.

We demonstrate the principles of this idea by applying it to the generation of simple figures known as Biomorphs (Dawkins 1986). The system is capable of learning which regions of Biomorph parameter space produce images that are displeasing to a particular user and can generate diverse candidate solutions of high quality. Furthermore, the solutions produced by combining the profiles of a small group of users are likely to be pleasing to all of the group members.

†To appear in *Artificial Life VIII: Proceedings of the Eighth International Conference on the Simulation and Synthesis of Living Systems*

An Aesthetic Immune System

An aesthetic immune system is a filter between users and a "stream" of art and eliminates undesired art before it can reach the user. Like the adaptive immune system, it "learns" from experience. If a user labels a work of art as "bad," the system will remember this and prevent this work from being shown to the user in the future. It would not be practical for the user to evaluate every item manually, so the algorithm must generalize. If the user dislikes a certain item, it is likely that the user will dislike similar items as well.

Aesthetic space is a framework for grouping "similar" works of art. In this space, the distance between the locations of two works of art is proportional to their similarity. Thus, similar works are close to each other while dissimilar works are far apart. For example, oil paintings of the Renaissance could occupy points in one part of the space, the paintings of the Enlightenment era might not be too far away, while the work of Abstract Expressionists would be quite distant from both of these. Clearly this space is impossible to construct accurately due to the subjective nature of similarity in the realm of aesthetics. However, when generating art algorithmically, we will assume that for many systems similar parameters produce aesthetically similar images.

If a user rejects one work, he or she will probably dislike most of its neighbors in aesthetic space. Therefore the system creates a *negative detector* that covers a small portion of space surrounding each rejected work. The detectors will behave like lymphocytes in the immune system, recognizing anything that is too similar to their pre-programmed targets. If a work is not rejected by any of the detectors, it is allowed to "survive." Therefore, the user should never be presented a new work that is similar to one explicitly rejected in the past.

The works of art that can pass through the system are not similar to previously rejected ones, but they are not guaranteed to be "good." If the user rejects one of these, the system will form a new detector for it, which will prevent similar "bad" works from being seen in the future. The untrained system with no detectors gradually constructs a profile of the user's tastes as it sees examples of art that the user rejects. Eventually, the accumulated detectors will form a model of the user's preferences and shield the user from a broad range of undesirable art.

To use an aesthetic immune system to help a user explore a wide variety of potentially good solutions, a random number generator can supply a stream of random candidate solutions to it. The system would filter out candidates that are likely to be bad, and when fully trained (if this is possible), it is a perpetual generator of novel but satisfactory art. Note that it does not attempt to find optimal solutions, so other techniques may be used to refine solutions presented to the user.

An important consequence of negative detection is that detectors can be independently generated by users with different judgment criteria. The areas that are not excluded by the superset of their negative detectors can define the regions of "consensus" solutions, or solutions that are satisfactory to everyone. Many users can generate "bad art" detectors independently, and when the detector sets of many users are combined, the system will generate works that none of the users will dislike. One can think of each user's detector set acting as a sieve, blocking the passage of art that is disagreeable to that user. The solutions that pass through all the sieves are the consensus solutions. This method could potentially satisfy the aesthetic demands of multiple users efficiently and accurately, even if the users are specifying their preferences asynchronously and incrementally. The quality of the consensus solutions depends on the domain and the heterogeneity of the users. If the users have conflicting interests, consensus solutions might not exist.

Related work

Evolutionary art (evoart) usually relies on humans to provide the feedback needed by the computer to guide the "evolution" of a work of computer generated art. Typically, the user iteratively refines a work by selecting a subset of favorite works out of a small set, which are variants or combinations of the user's favorites from the previous time step. After many iterations, the quality of the works can improve. This process of "evolving" art is sometimes called *aesthetic selection.* Dawkins was the first to implement evolutionary art on a computer (Dawkins 1986). In his system, the user explores the space of images called "Biomorphs" in the iterative manner described above.

The aesthetic immune system approach has several advantages over the evolutionary approach. First, evoart systems usually require a user to compare many members of a population at once to pick a favorite, while an immunologically-based system does not. For some systems, such as the generation of musical phrases, it is difficult to evaluate many individuals at once (Nelson 1993). Second, the evolutionary process converges to single works of art, while the immunological approach evaluates the whole of parameter space and can be a source of perpetual novelty. Third, in order to produce innovative works, evolutionary systems often combine works of art (Sims 1991). The programmer is faced with the difficult task of defining operators to mix the "genotypes" of works in a manner such that the offspring's "phenotype" exhibits desirable traits of its parents. In some cases, the results are unpredictable and displeasing, suggesting that more complex combination rules are required. A trained aesthetic immune system can generate a very large range of aesthetically pleasing art without the use of a difficult-to-define art combination operator. This may be crucial in collaborative evoart, in which many users cooperate to generate works of art. Simply combining the favorite works of different users might yield unsatisfactory results. If a fan of Salvador Dalí's surrealist canvases were to meet a fan of Norman Rockwell's sentimental and more realistic style, a painting that combined the two painters' tendencies would probably be unsatisfying to both. A better solution might be to introduce both to something completely different, like the paintings of Paul Klee. This is the sort of solution an aesthetic immune system would be capable of proposing.

Application to Biomorphs

We applied the immunological framework to the exploration of Dawkins' Biomorphs (Dawkins 1986). A Biomorph is a recursively drawn figure defined by a nine-digit "genotype." The digits, all but one of which are allowed to take the values between -9 and 9 inclusively, are parameters that affect the appearance of the final image, or "phenotype." The final digit of the genotype, which defines the level of recursion used in drawing the Biomorph, is restricted to values between 2 and 9. The following sections describe how to implement an aesthetic immune system to help users find aesthetically pleasing regions of the Biomorph genotype space.

Defining the aesthetic distance

The first step is to define the distance between any two solutions. Fortunately, small changes in Biomorph parameters (the genotype) usually result in small phenotypic changes, so we were able to use genotype space as a proxy for the true aesthetic space. That is, Biomorphs with similar genotypes are assumed to look similar, while the appearance of Biomorphs with very different genotypes is assumed uncorrelated. For convenience, we define the distance in genotype space to be the maximum arithmetic difference between any two corresponding digits of the two genotypes, which is the Minkowski metric $L_\infty(a,b) = [\sum_i (|a_i - b_i|^\infty)]^{1/\infty}$, where a and b are the strings being compared and a_i and b_i are the ith element of a and b. For example, the distance between the two genotypes (0,0,0,0,1,8,1,1,2) and (7,3,6,2,1,1,3,3,2) is 7 because that is the maximum difference between the corresponding digits of the two strings (at both positions

1 and 6). Other distance measures are possible, such as the sum of the differences between corresponding positions in the genotype.

For some domains, one would not want to generate new works, such as with a collection of already existing work. In such cases, the immunological approach could be used to prevent the retrieval of entries that are too similar to previously rejected works. There is a large body of research on the problem of quantifying similarity between data, including content-based classification, metadata-based classification, and collaborative filtering techniques.

Defining the detectors

A detector should be large enough so that the user does not need to define "too many" of them but not so large that they eliminate as many good solutions as bad ones. We will define a detector of radius r to cover all genotypes within distance r from the center of the detector. The neighborhood size of a detector is the number of distinct locations in space that it covers. The detector neighborhood should cover a "large enough" portion of the total parameter space such that a reasonable number of detectors could cover a large portion of the total genotype space. It may be fruitful to augment the immunological approach with pattern classification techniques to more efficiently cover parameter space.

For the Biomorph user experiment, described below, we used large detectors in order to minimize the amount of time it took the subjects to cover a significant portion of space. We chose a detector size such that the average user could create enough detectors to fill about half of Biomorph parameter space in under ten minutes. The large detector size affects the results quantitatively, but hopefully not qualitatively. It is probably impossible to guarantee that no good genotype will be censored by a detector, but we would like to minimize this possibility. If the transition between high and low quality solutions is not too abrupt, there could be regions of expendable intermediate-quality solutions between them in parameter space. If the user is instructed to not reject solutions of low but acceptable quality, the detectors are less likely to interfere with good solutions.

We set the detector radius to be 6 for all parameters except for the last one, which we set to 2. Small changes in the last parameter greatly affect the appearance of Biomorphs, so a smaller radius is appropriate. Figure 1 shows a random Biomorph and various others that fall within its detector radius. Based on this and several other examples, we concluded that Biomorphs that are within the range of a single detector are qualitatively similar and assume that if a genotype yields an undesirable phenotype, then most genotypes within its detector are likely to be undesirable as well.

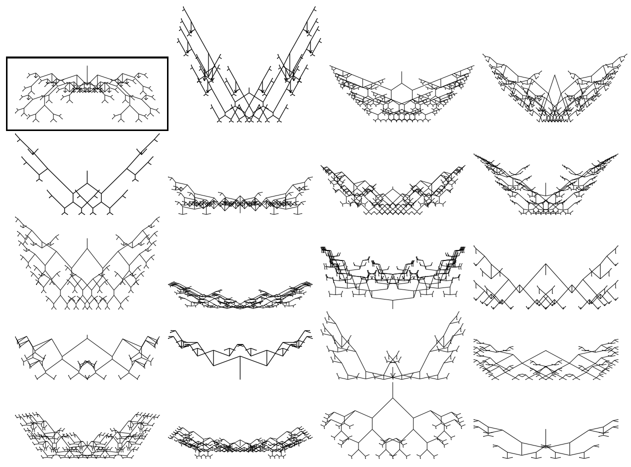

Figure 1: A sample from the range of Biomorphs covered by a single detector. The boxed figure in the upper-left is the detector's center, which is (3,6,6,-5,3,-2,-6,-5,8).

Experiment

Seven volunteers participated in a user study to determine the feasibility of the aesthetic immune system. All used a program to create their own Biomorph detector sets and were retested several days later to determine the effectiveness of system's detectors. To create a Biomorph detector set, a computer program displayed a series of randomly generated Biomorphs. For each Biomorph shown, the subject was given the option to reject it, creating a detector based on its genotype, or to accept it, which does not change the state of the detector set. The detectors were put to use as they were created to filter out new Biomorphs similar to the rejected ones. Thus, the aesthetic quality of images was expected to improve during a session because the increasing number of detectors could prevent more regions of "bad" genotypes from being displayed. The subjects were asked to rate Biomorphs until about 50% of the parameter space was covered by detectors. To test the effectiveness of the detectors, the subjects were asked to evaluate a series of 150 Biomorphs from three sets: random Biomorphs, Biomorphs that survived censoring by the subject's own detectors, and Biomorphs that survived the censoring by the superset of many subjects' detectors. The number of rejected Biomorphs from each set was recorded, and we assume user satisfaction to be inversely correlated with the number of rejections.

This test was performed using all seven subjects and the superset of all subjects' detectors to determine the efficacy of the consensus solutions. It was repeated using only three of the subjects and their detectors to determine the effects of group size on the quality of consensus solutions. The consensus solutions shown in the latter test were thus censored using only three sets of detectors instead of seven.

a)

	1	2	3	4	5	6	7
none	52%	72%	50%	30%	58%	36%	54%
self	38%	48%	44%	32%	66%	30%	44%
all	46%	62%	38%	48%	64%	38%	60%

b)

	1	2	3
none	62%	52%	50%
self	56%	48%	28%
all	44%	42%	30%

Table 1: The user study results. The entries represent the percentage of Biomorph images rejected by a user for a particular detector set. Each column represents the data from one subject. The detector sets, indicated by the labels "none", "self" and "all", are no detectors, the subject's own detectors, and the superset of all subjects' detectors respectively. The results from testing all seven users are summarized in a), and the the results from the subset of three are shown in b).

Results

The results of the experiments are summarized in Table 1. The system appeared to learn the Biomorph preferences of individual users to a statistically significant degree. For the testing involving all seven users, the rejection rate for the random Biomorphs was higher than that for Biomorphs that were filtered using a subject's own detectors ($>$ 95% significance using the Wilcoxon signed-rank test). For single users, the system clearly reduced the number of unacceptable solutions, thus improving the average quality of solutions presented to an individual. We believe that the system's performance could be improved dramatically for motivated users who are willing to invest more time to build their personal detector sets (e.g. for music, movies, or books that they like.) Using smaller detectors and covering a larger portion of the parameter space would likely increase the accuracy of user preferences. More importantly, a more accurate aesthetic distance measure needs to be defined.

Combining user preferences worked well for the set of three users but not for all seven. For the testing involving the seven users, the difference between rejection rates of random Biomorphs and those censored by the superset of all seven subjects' detectors is not statistically significant. For the testing involving only three of the seven subjects, the rejection rates for the group-selected Biomorphs were approximately the same or lower than the ones selected using their own detectors. The consensus solutions for three users appeared to be of equal or greater quality than the solutions generated for the individuals alone, while the solutions for all seven users appeared to be of low quality. The results suggest that it is hard to satisfy a large number of users and that consensus solutions may not even exist for groups with members with different preferences. The number of users that an aesthetic immune system can accommodate is likely dependent on the particular problem domain and the similarity of the users. If the users have conflicting preferences or if the problem domain is highly subjective, then consensus solutions would be hard to find. However, if user preferences are orthogonal to one another, then a large range of good consensus solutions could be discovered using this method. Inaccuracies in the user profiles can also lower the average quality of solutions found. If the profiles for all users were perfect and complete, solutions that are not censored by any profile would satisfy everyone. However, the Biomorph profiles were approximate and incomplete, so combining profiles did not always work.

Conclusions

We have presented an algorithm for identifying regions of interest in a parameter space using the subjective judgments of multiple individuals. It remains to be seen if the consensus solutions are more likely to be novel and innovative or bland and inoffensive. It would be easy to retrofit many existing evoart algorithms to use the immunological approach. We believe that this will reveal interesting parts of the parameter spaces that were not seen before due to the inefficiency of mutational search and the local aesthetic fitness optima that can stall evolution. The search for these regions could be accelerated by having multiple users create a common pool of detectors. If the set of cooperating users is large and representative of the tastes of the general public, the final result might be a system that generates art that many can appreciate.

Acknowledgments

We thank the participants in the user study. The authors gratefully acknowledge the support of the NSF (grants CDA-9503064 and ANIR-9986555), the ONR (grant N00014-99-1-0417), DARPA (grant AGR F30602-00-2-0584), the Intel Corporation, and the Santa Fe Institute.

References

Chao, D. L., and Forrest, S. 2002. Information immune systems. In *Proceedings of the First International Conference on Artificial Immune Systems.* (in press).

Dawkins, R. 1986. *The Blind Watchmaker.* Harlow, UK: Longman Scientific and Technical.

Nelson, G. L. 1993. Sonomorphs: An application of genetic algorithms to the growth and development of musical organisms. In *Proceedings of the Fourth Biennial Art & Technology Symposium*, 155–169.

Sims, K. 1991. Artificial evolution for computer graphics. *Computer Graphics* 25(4):319–328.