Objet Trouvé, Holism, and Morphogenesis in Interactive Evolution

Ron W. Noel (noelr1@rcn.com)
WCSU, Department of Psychology, 181 White Street
Danbury, CT 06810

Sylvia Acchione-Noel (sylvia.acchione-noel@corporate.ge.com)
General Electric, JF Welch Leadership Development Center, Old Albany Post Road
Ossining, NY 10562

Abstract

Evolutionary systems are conceptualized as having four transfer functions between the two state spaces of genotype and phenotype. The four transfer functions are epigenesis, survival, mate selection, and genetic recombination. The treatment of these transfer functions is uneven at best. In particular, some complain that epigenesis, the formation of an entity from the original undifferentiated cell mass into tissues, is treated in a too simplistic manner to allow for system flexibility, or creativity. This paper reports on an interactive image evolving system that mimics the morphogenesis processes in epigenesis. System description, results, and theoretical implications are discussed.

Introduction

Interactive evolutionary systems seek to interface evolutionary programming to human preference in order to create systems capable of evolving artifacts that require a human expertise that hasn't yet succumb to computation. A common area for this endeavor is the evolving of art, particularly image. The interfacing of human ability with machine computation requires resolving difficult issues in the arts, humanities, and sciences. Further, progress in the design of interactive evolutionary systems allows a glimpse into how very human abilities such as intuition, projection, and holistic perception interplay with the mechanics of machine computation. This paper reports on one such interactive evolutionary system that seeks to combine human perception with the genetic algorithm to evolve small holistic images.

Humans lack the tremendous numerical computational speed of computers; yet they can process information holistically in an automatic, rapid, and natural manner. Machines possess tremendous computational capabilities; yet no algorithm exists to perform holistic processes as well as humans do. Ideally, a good interactive system would integrate the best human cognitive qualities with machine computational capabilities enabling the resultant hybrid system to outperform either of the two components alone. Evolutionary computation as an algorithm is well suited for the creation of an interactive imaging system. However, problems exist in implementation: How can evolutionary computation best support the holistic processes of human cognition? To answer this question requires an understanding of current theory regarding human holistic processes.

Psychological Issues

A well-known area in cognitive research that studies holistic processes is the recognition of objects and, in particular, the recognition of faces. Different perceptual encoding and representational processes have traditionally differentiated theories regarding the recognition of objects as compared to faces. However, the functional separation of these processes under all conditions of object recognition remains unclear (Bruce & Humphreys, 1994). Much of basic object recognition theory has been based on the decomposition of parts and the analysis of edge features (Marr & Nishihara, 1978; Biederman, 1987; Ullman, 1989). On the other hand, face recognition theory has been based on more holistic processes which utilize surface characteristics such as texture, color, and shading (Price & Humphreys, 1989). Some research suggests that the distinctions between object and face recognition begin to fade when one examines the object recognition processes of experts, who may utilize holistic processes similar to those found in face recognition (Diamond & Carey, 1986; Rhodes & McLean, 1990).

The theory regarding holistic processing of faces can be separated into stronger and weaker stances (Bruce & Humphreys, 1994). Under the weaker stance, features may interact with each other through configural processes to form emergent properties or "second-order relational features" (Diamond & Carey, 1986). Under the stronger stance, face recognition is completely holistic; that is, its representation is non-decomposable in that no explicit description of features exists outside the context of the face (Tanaka & Farah, 1993). These stances provide two ways of approaching the development of an interactive system to support the holistic development of images: (1) A system which manipulates context-free features towards configuration, or (2) a system which develops the configuration of the image first, followed by more detailed development of features within the established context.

We sought to design a recognition-driven system of the latter type, which would support the purely holistic development of images, including faces and objects. This system would function in a feature-free space to provide a non-decomposable representation of images. For instance, a human may perceive or project a cloud as containing the image of a face, yet a cloud contains no explicit representation of the eyes, nose, or mouth. Such features, say a nose, would only be perceived as a nose within the context of the cloud's facial image. This type of perception or projection of a natural texture is called *objet trové*, and is thought by some (Gombrich, 1960) to be paramount to the perception of art. Further, a cloud is not limited to faces; it might contain other animals, or objects. The cloud merely contains randomly distributed textures which humans can organize perceptually. Just as a cloud supports holistic recognition of images, an interactive evolutionary system could encourage recognition based upon context-dependent rather than context-free properties. We intended to provide a mechanism by which a cloud-like representation could enter into "cumulative selection", in a manner not unlike the wishful thinking of Dawkins (1987).

System Issues

The system presented in this paper can be distinguished from other work in the interactive evolution of images (Dawkins, 1987; Baker & Seltzer, 1994; Sims, 1991; Caldwell & Johnston, 1991; Johnston & Caldwell, 1997; Todd & Latham, 1992). Many interactive evolutionary systems (Dawkins, 1987; Baker & Seltzer, 1994; Sims, 1991; Todd & Latham, 1992) use aesthetic preference to determine the fitness of images that are composed of context-free features. Under conditions of aesthetic preference, the user evolves images opportunistically. These systems do not easily support evolution towards an *a priori* goal. Baker and Seltzer (1994) opportunistically evolved butterflies from randomly generated lines, but when they intentionally evolved a general "face-like" image, they could do so only with difficulty. Further, previous systems sometimes required input images to enable the evolution of faces. Sims (1991) as well as Baker and Seltzer (1994) modified facial images after providing input images of human faces, and Johnston and Caldwell (Caldwell & Johnston, 1991; Johnston & Caldwell, 1997) provided input images of features to evolve configured faces.

The Johnston and Caldwell system (Caldwell & Johnston, 1991; Johnston & Caldwell, 1997) is most similar in purpose to our system in that they used human recognition to evolve images of criminal suspects. Their “Faceprints” system allowed more goal-directed behavior within interactive evolution than previously achieved, and they developed a system that encouraged holistic processes by presenting configured faces from the start. However, their system differs from ours in that they took the weaker theoretical stance towards holistic processing by providing input images of context-free features and then placing them in a randomly generated configuration for further evolution.

The “Faceprints” system represents an approach which is common to evolutionary computation; that is, the majority of evolutionary computation is based on parameterized models which predefine features and pre-encode dimensions upon which the features can vary. A key component of evolutionary computation is the mapping between the genotype and phenotype representations. The genotype representation consists of a string of characters, usually binary, that are used as genetic codes in the reproductive process. The phenotype representation consists of a description of an organism that can be evaluated for fitness and selected for reproduction. The linkage between the genotype and phenotype representations is accomplished by a mathematical mapping that uses a parameterized model. For instance, to evolve rectangles, one could create a formula with the two parameters of height and width that would scale suitable binary numbers to a rectangle of a certain height and width. The binary numbers would form the genotype, and the resultant rectangles would form the phenotype.

There are many problems associated with approaches based on a parameterized model (Hofstadter, 1982). The main problem for creating images is that parameterized models constrain the phenotype representation. A model for rectangles can never create a circle. We might try to escape the problem by adding a selector parameter that would dictate the geometric shape to use. For instance, in a rectangle model, if we wanted to represent circles also, we could add a selector parameter that would indicate whether to implement a rectangle model or a circle model. The repair works, except that the addition leads to a discrete selector parameter function and potentially requires an infinite number of models to represent all objects. Instead, we seek to create a system that avoids the predefinition of features and the mapping of genotype to phenotype.

Image Elicitation System

Instead of using a feature-based space, we created a frequency-based space based on pixel representation. The pixel space representation affects the resolution of the images, but forces no predefined features upon the images themselves. The space is based upon atomic or molecular representation, similar to the notions of atomic or molecular decomposition by Fourier analysis or wavelets (Meyer, 1993). As championed by the pointillists, small points of just a few colors can be used to create the psychological impression of any form and any color. Atomic representation works at the sub-feature level and allows the generation of features along with their configuration. The representation is not constrained by features and encodes a dog, a tree, or a car as easily as a face. For instance, one could create a pixel space of 25-by-25 pixels (625 total units) with each pixel being any of eight colors (three bits of information.) Such a small pixel space has the informational potential to create an enormous number of images, as many as 2^{1875}. The number of possible images is so large that there exists no real constraints on the variety of forms that may be represented; rather, the model constrains the resolution of the image. In terms of resolution, the space cannot represent objects that require more than 12.5 lines of resolution (each line requires at least one 'on' and one 'off' pixel) in the vertical or horizontal axis. However, increasing the number of pixels and decreasing the pixels' size can reduce the impact of the constraint.

System implementation required addressing additional issues in the method of reproduction and mutation function. First, usually, simple cross-over points are used as the method of reproduction, but such a linear system is inappropriate for a two-dimensional space. Instead, we increased the number of cross-over points until the reproductive system considered a cross-over point at every allele. Such a system of uniform crossovers was implemented by randomly selecting between the genes of the two parents with equal probability. Although some researchers consider uniform crossovers to be deleterious to evolutionary computation (Fogel, 1995), others have found them to be useful (Syswerda, 1989). Secondly, if one uses a mutation function that chooses among all possible genes with equal probability for an allele, the mutation function will eventually return the image to a random state. Instead, we limited the mutation function to the gene values of neighboring pixels, causing smaller changes and greater adaptability.

Each generation has a population of fifty images of which the human selects ten images for use as the parents for the next generation. The resulting image elicitation system consists of a comma plus system since the parents are available for selection in the next generation so that each generation after the first is made of parents plus their offspring (Heitköttere & Beasley, 2002). The genotype representation is an array of alleles that has the same size as the pixel representation (25x25 pixels). Each allele is a character that corresponds to one of the possible colors (or genes) for the pixel. Reproduction creates the offspring genotype by randomly and uniformly selecting between the genes of two randomly selected parents at each allele site.

Results

For purposes of this paper, and given our emphasis on holistic processes, we chose a face as the image to be elicited. The first author began with the specific goal of "elicit Abraham Lincoln" and elicited an image of Abraham Lincoln using four levels of gray (figure 1). The image represents the results of image elicitation after 245 generations. The image was originally generated on a SVGA monitor using a black background. A human's ability to recognize Abraham Lincoln is very dependent on the spatial frequency of the image. In other words, viewing figure 1 at too close or too far of a distance reduces the perceptual quality of the image. Figure 2 displays the evolution of the stochastic prototype of Abraham Lincoln. The matrix represents every fifth generation of Abraham Lincoln's image up to generation 245. The matrix should be scanned from left to right and from top to bottom. Each image is a stochastic prototype created by randomly selecting and copying a gene from one of the ten parents into the corresponding allele of the prototype until each allele of the prototype is created. The sampling function is a uniform, random distribution over the parents. As a result, the composite prototype is similar to all of its parents and evokes the average recognition of its parents.

Theoretical Impact

The process in image elicitation is best described in terms of co-evolution or holistic evolution. This description runs contrary to mainstream thought on how evolutionary computation works. There are currently two ideas on how convergence happens in evolutionary computation: the Building Block Hypothesis by Goldberg (1989) and the Schema Theorem by Holland (1992). We argue that both of the hypotheses are feature analytic and are insufficient to explain what is happening in image elicitation.

The Building Block Hypothesis suggests that the convergence process in evolutionary computation is based upon building blocks or small groups of characters whose introduction into any genotype representation will likely increase the fitness of the phenotype representation. Goldberg suggests that the genetic computation first finds as many of these building blocks as possible, and later in evolution, the building blocks are combined together to give the highest fitness. For instance, a series of 1's in the genotype might give rise to an eyebrow in the image. The presence of an eyebrow in any picture of a face should increase the image quality, and, therefore, the fitness.

The Schema Theorem suggests that the ongoing process in evolutionary computation is implicit parallelism caused by schema processing. Schemata are defined as patterns of characters in the genotype representation that may include "don't care" states. A schema can be specified by a genotype representation in which each gene contains a 1 for "on", 0 for "off", or X for "don't care". In a sense, a schema is a relaxed building block in that it relaxes how tightly clustered the group of "care" genes are. Each genotype representation can contain a large number of schemata. This leads to the implicit parallelism and speeds up search.

The basis for our criticism of the current theories lies in their assumption that one can analyze the genotype while disregarding the phenotype. It also requires one to accept that all intermediate representations (patterns in the genotype that are tried and not kept) are coincidental to the process. In such a view one need only look back from the evolved solution and trace the heritage of its genes. In both theories, the implied process is analytical.

Image elicitation challenges these theories in terms of process and representation. Image elicitation relies on multiple-gene (holistic) representation as opposed to variable-encoded (feature-based) representation. In image elicitation, the image exists in the phenotype and in the perception of the observer, whereas, in other evolutionary computation, the image description exists in the genotype. Our system allows polygeny and pleiotropy, whereas current theory is based on a direct mapping from the genotype to the phenotype and no separate mapping backwards from the phenotype to the genotype. We can illustrate our theoretical differences through what we call "the gray argument" for holistic processes.

Consider the evolution of a medium gray. Mapping from the gray phenotype back to the genotype reveals two optimal representations as shown by *a* and *b* in figure 3. Using binary representation, where 1 equals an "on" pixel, 0 equals an "off" pixel, and X equals "don't care", one finds that 1010101... is one representation *(a)* and it's complement 0101010... is the other *(b).* Strangely, one can breed together these complimentary representations and the offspring *c* and *d* will still appear grayish in the phenotype, regardless of genotype of the parents, or type of crossover function used. Grays *c* and *d* result from the use of simple and uniform crossover respectively. How can one describe the process of evolving grays in terms of building blocks or schemata when the phenotype being selected does not depend upon any particular gene being a 1 or 0 or X? The gray phenotype can only be described in the genotype as a fairly uniform distribution of 1's and 0's. Given that the difficulties in phenotype-to-genotype mapping of grays extend to all images, the Building Block Hypothesis and the Schemata Theorem are insufficient to describe an image's evolution. In fact, the gray argument is problematic for many current approaches used to understand, or represent phenotype-to-genotype mapping.

Conclusions

Image elicitation promotes wholeness or a lack of distinct features. If one observes the evolving of Lincoln's image one will notice that the features all co-evolve together. At no time does the process treat the nose differently from the eye. The process is so tightly interwoven that to distinguish the nose from the rest of the face constitutes a false distinction; the nose gets it description from itself and the context provided by the rest of the image.

Image elicitation affords high-order interactions. The placement, sizing, shading, and coloring of an image that bears strong resemblance to the original (e.g., a face) can only be described as highly interactive. The placement, size, etc. of the nose is dependent upon all the other features of the face, for if the nose is anywhere but the right place in relation to other features the image would have no resemblance. High-order interactions are a problem for analytical processes, but not this method. It evolves a complex stimulus within a large information space while maintaining a small population size and a reasonable number of generations.

Image elicitation appears to promote holistic rather than analytical processes. It begins with the grossest of detail and ends with the finest of detail. The elicitation process uses intermediate approximations as placeholders for features and as a means of resolution building. This process of organizing an undifferentiated representation into finer and finer regions, or features appears to mimic the process of morphogenesis found in biological reproduction. We argue that features, which by their nature are fine detail and not available until the end of convergence, cannot be used to explain the process. Such findings are problematic for the building block and schemata theories that are currently used to explain the processes in evolutionary computation.

Summary

The present image elicitation system provides a new technique for integrating the best qualities of human and machine capabilities to create images. Neither system could produce theses images alone. Machines lack the perceptual and memory skills, and humans lack the computational energy to evolve an image. The results show that current theories of evolutionary computation are insufficient to explain the convergence of the images in the absence of a feature-based parameterized space.

The technique of image elicitation allows humans to use their perceptual and cognitive systems to organize visual noise into the objects of their memories. This process of literally pulling an image out of chaos will affect our understanding of intelligent systems and future investigations across many disciplines. Image elicitation will be useful in studying machine intelligence, as well as in studying top-down processes in interactive intelligent systems. The system provides a means for humans to experience how evolutionary computation works by directly immersing themselves in the process. And, it provides cognitive researchers with a means of studying holism in human recognition.

Figures

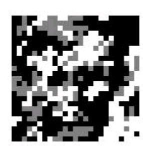

Figure 1: A stochastic prototype of Abraham Lincoln at generation 245.

Figure 2: The evolution of an image of Abraham Lincoln, showing every fifth generation up to generation 24

(a.) (b.)

(c.) (d.)

Figure 3: The gray argument for multiple-gene (holistic) representation. Grays *a* and *b* have complimentary phenotype. Grays *c* and *d* result from the use of simple or uniform crossover, respectively.

References

Baker, E. & Seltzer, M. (1994). Evolving line drawings. *Graphics Interface '94 Proceedings*.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*.

Bruce, V. & Humphreys, G. (1994) Recognizing objects and faces. *Visual Cognition*.

Caldwell, C. & Johnston, V. (1991). Tracking a Criminal Suspect through "Face-Space" with a Genetic Algorithm. *Proceedings of the Fourth International Conference on Genetic Algorithms*, Morgan Kaufman.

Dawkins, R. (1987). *The blind watchmaker*. Longman Scientific and Technical, Essex, UK.

Diamond, R. & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*.

Fogel, D. (1995). *Evolutionary computation: Toward a new philosophy of machine intelligence*. IEEE Press, New York.

Goldberg, D. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley.

Gombrich, E. (1960). Art and illusion: A study in the psychology of pictorial representation. Princeton University Press, Princeton.

Heitköttere, J. & Beasley, D. (2002). *The hitch-hiker's guide to evolutionary computation*. www.cs.bham.ac.uk/Mirrors/ftp.de.uu.net/EC/clife/www/ - 7k - 04 Feb 2002 .

Hofstadter, D. (1982). Metafont, metamathematics, and metaphysics: Comments on Donald Knuth's article 'The concept of a meta-font'. *Visible Language*.

Holland, L. (1992). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. Cambridge, MA: MIT Press/Bradford Books.

Marr, D. & Nishihara, N. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London*.

Meyer, Y. (1993). *Wavelets: Algorithms and applications*. (Translated and revised by R. Ryan). Society for Industrial and Applied Mathematics, Philadelphia.

Price, P. & Humphreys, G. (1989). The effects of surface detail on object categorization and naming. *Quarterly Journal of Experimental Psychology*.

Rhodes, G. & McLean, I. (1990). Distinctiveness and expertise effects with homogenous stimuli: Towards a model of configural coding. *Perception*.

Sims, K. (1991). Artificial evolution for computer graphics. *Computer Graphics*.

Syswerda, G. (1989). Uniform crossover in genetic algorithms. *Proceedings of the Third International Conference on Genetic Algorithms*. J. D. Shaffer (Eds.), Morgan Kaufmann Publishers, Los Altos, CA.

Tanaka, J. & Farah, M. (1993). Parts and wholes in face recognition. Quarterly Journal of Experimental Psychology.

Todd. S. & Latham, W. (1992). *Evolutionary art as computers*, Academic Press: Harcourt Brace Jovanovich, London.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition*.