Tableau Machine: A Creative Alien Presence

Adam Smith[†], Mario Romero[‡], Zachary Pousman[‡], and Michael Mateas[†]

† University of California Santa Cruz
Department of Computer Science
1156 High St., Santa Cruz, CA, USA
{amsmith, michaelm}@cs.ucsc.edu

‡ Georgia Institute of Technology
College of Computing
85 5th St. NW, Atlanta, GA, USA
{zach, mromero}@cs.ucsc.edu

Abstract

We present the design of *Tableau Machine* (TM), an AI-based, interactive, visual art generator for shared living spaces. TM is an instance of what we call "alien presence": an ambient, non-human, embodied, intelligent agent. From overhead video in key public spaces, TM interprets its environment, including its human audience, and expresses its interpretation by displaying a sequence of abstract images of its own design. This paper is a case study in the design of an art generator with deep and long-term connections to its physical and social environment.

Introduction

Tableau Machine (TM) is an interactive, generative art installation designed for deployment in shared living spaces such a homes or offices (Romero, Pousman, and Mateas 2007). Figure 1 shows a scene from a typical installation. The system has a bright, color display mounted in a prominent area that displays a smoothly-fading sequence of abstract, visual compositions (style inspired by the Russian Constructivists). The display, together with an attached printer, conveys an expression of the system's interpretation of the living space as detected by its sensors. TM includes several video cameras mounted throughout the living space acting as very precise motion detectors. Finally, connecting the physical components together, a standard PC runs our software controlling the complex and adaptive mappings from motion data to displayed compositions. The system is complete and has been installed in homes where it shares the living space with a human audience for approximately six weeks.

At the software level, TM uses basic image processing and a semantically-motivated data abstraction technique to reason about its environment. Next, TM creates visual outputs using a set of design grammars. The final images are selected on the basis of visual properties described by abstract shape-generation rules in the grammars as well as emergent visual properties detected by a pixel-level image analysis.

TM was designed as an alien presence (AP). An AP is a non-anthropomorphic social entity embodied by a computer, imbued with AI-based perception, interpretation, and artistic output (Romero and Mateas 2005). The main purpose served by an AP system is to create a long-term interpretive experience for its audience. In particular, this experience should engage human meaning-making behavior as audience members build theories to describe the AP's behavior. In this paper we will focus on the following functional definition of AP: a non-human system that expresses its interpretation of its environment over the course of a long-term interaction.

Figure 1: A scene from a typical installation of TM showing the large display and several cameras.

We cannot, with rigor, claim that TM is truly creative. However, as a result of both our design and some technical limitations, TM addresses several issues that frequently arise in machine-creativity discussions: a generate-and-test framework, a search for novelty and value, and exploration of a conceptual space.

Our main contribution is a complete, non-trivial, generative art system, an alien artist in the home, which intentionally produces complex images in direct, meaningful response to the system's observations of a human audience. Our second contribution is a method for gaining expressive control over the output of design grammars by using image analysis.

Related Work

TM is an interdisciplinary work; in this section we will focus on the specific connection between TM and the areas of generative art systems and the evaluation of machine creativity.

Generative Art

AARON is a prominent generative art system developed by Harold Cohen (1995). AARON uses a large database of painting rules to generate images of people and plants. AARON uses no human input (other than Cohen's original programming) and has a unique style of its own. Interestingly, despite being billed as one of the best examples of creative generative art systems, Cohen (1999) himself refuses to attribute creativity to the system.

In contrast to AARON, the generative art system NEvAr (Neuro Evolutionary Art) (Machado and Cardoso 1997) does not have a fixed procedure for generating images. Instead it uses genetic search to evolve small programs composed of a tree of low-level mathematical operations that produce abstract output images, often harnessing human feedback on sample outputs to tune an artificial neural network that guides the evolutionary process. NEvAr's authors refer to the system as a constructed artist with an internal, adaptive sense of visual aesthetic value.

There are numerous generative art systems but these two give sufficient comparative context to TM. Like AARON, TM uses pre-programmed rule sets for producing its output where the specific rules do not represent a general theory for machine creativity. Like NEvAr, TM's outputs are very abstract in nature and are not purely the result of a random number generator, but instead are the result of programmatic reasoning, by the system, about the pixel-level appearance of images with respect to a model of perception. Unlike other generative art systems, TM's design, as a complete agent sensing the world, encompasses more than image generation. Furthermore, human interactions with the system take place on a much longer time scale than other generative art systems.

Machine Creativity

In the interest of brevity, we only touch on some high-level issues in machine creativity. In "Evaluating Machine Creativity", Pease et al. (2001) assert that while there is no clear agreement on what it means for something to be creative, repeated concerns in the literature define a creativity space (supporting questions of the form "Where does *x* lie in creativity space?"). A central idea is that novelty and value are necessary conditions for creativity.

The presence or amount of novelty and value can be measured from a variety of perspectives by several means. The most common method discussed for procedurally defining novelty and value involve a generate-and-test process. Ideally, the generation phase synthesizes novel artifacts and the testing process selects only those artifacts that are valuable, resulting in creative outputs. Using these outputs as inputs for another round of generate-and-test is commonly called the "central loop of creativity" and illustrates the requirement that something is always produced in a creative process (though it may be abstract). TM uses a generate-and-test process, and is designed so that its outputs might be perceived as novel and valuable. However, organized as a pipeline, TM never consumes any of its outputs and thus cannot use them for the basis for new generation.

Many creative processes can be seen as harnessing a source of randomness for use in a deterministic process. TM makes use of software-generated random numbers in several ways; however its nondeterministic, *real-world* inputs play a far more important role in its behavior.

Finally, internal reflection, metacognition, or thinking at the meta-level about the process being executed is sometimes considered a necessary condition for creativity (Buchanan 2001). Buchanan argues that, so far, attempts at building creative computer programs fall short of achieving their goal because "(1) they do not accumulate experience [regarding their own internal processes], and, thus, cannot reason about it; (2) they work within fixed frameworks, including fixed assumptions, methods, and criteria of success; and (3) lack the means to transfer concepts and method from one program to another." TM fares no better than other programs along this dimension of analysis.

Goals

As TM was originally motivated by research in expressive AI (Mateas 2001), human-computer interaction, and generative art, we have several goals for TM.

First, the system should be an unfamiliar presence with non-anthropomorphic agenthood. Non-anthropomorphic systems are not expected to "understand" the idiosyncrasies of human behavior, but may be expected to "understand" much more general physical or statistical patterns. The design flexibility that a non-anthropomorphic system affords its designers is of interest in expressive AI.

Next, the system should form an interpretation of its environment (specifically the human audience), and express this interpretation via its visible output. That is, the system should be a participating occupant of the living space to the level that its physical design allows. The way in which the system affords building explanations (correct or not) of the system's behavior is relevant to human-computer interaction.

Finally, the system's output should be visually interesting, both in the sense that it has aesthetic value and is relevant to the situation in which the output arises. The generative art concern here is with the size and ease of exploration of the generative space of outputs as well as the ability for the system's authors to expressively shape the space.

Tableau Machine as an Agent

The design of TM is driven by an intelligent agent perspective. Agents are embodied systems consisting of clearly defined input and output interfaces connected by a mapping called the agent function.

TM has several low-level inputs (percepts, means of sensing). The system is aware of the time of day, a live sequence of video frames showing wide-angle, overhead views of several locations in the living space, and a human-designed map of the physical space in which the system resides.

Next, TM's low-level outputs (effectors, means of taking action) are focused on the display. The agent function formally outputs a specific image along with a delay time (before selecting a new image), effectively forming an infinite animation.

Finally, the agent function is structured in terms of the AP-dictated interpretation and expression pipeline. Interpretation can be though of as a lossy compression process, a many-to-few mapping of the space of all inputs, to a smaller space of models that loosely explain the input. Alternatively, interpretation can be thought of as an abstraction process, describing properties of inputs while forgetting their concrete details. Expression can be thought of as partial expansion process, a few-to-many mapping between models and detailed output data. Similarly, we can think of expression as a de-abstraction or grounding process that makes concrete artifacts from general requirements. In this way, a particular model may be expressed in a variety of ways. In the following sections we will describe in detail the interpretation and expression processes in TM.

Interpretation

Focusing on the interpretation process in TM, we will describe how the formal inputs to the system get mapped down to models. Our goal in interpretation is not to simply squeeze bits out of the input stream, however. We designed the system to build representations of its input in a way that supports the kind of stories the audience might invent about how the system behaves. This constrains the space of interpretation processes to those that produce models that are simple enough for the system to meaningfully express. Figure 2 gives a preview of how our three-level model of the environment is assembled via interpretation.

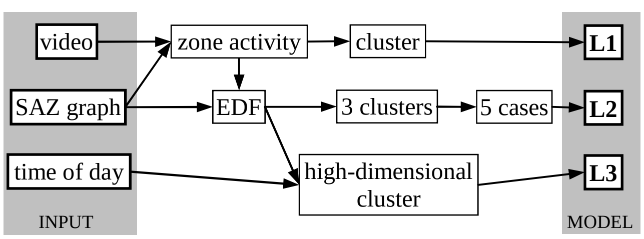


Figure 2: Building an interpreted model.

Semantic Activity Zones and EDF

The video frames streaming into the system from the cameras provide far more detail than we intend our alien to understand. We turn video frames into a measure relevant to the living space by exploiting an author-provided map that tells the system which areas of each image correspond to distinct zones of the living space. We call this map the semantic activity zone (SAZ) graph because zones are selected to represent spaces that, if actively occupied, might imply that a person was performing some distinguishable activity. Seats on a couch in the living room or the space in front of the dishwasher are examples of such SAZs. We do not tell the system what each SAZ "means"; SAZs only provide image-space distinctions. Each node in the SAZ graph defines a region in the camera views, while weighted edges describe the connectivity and distance between SAZs ("distance" is a measure of how easy it is for a human to move from one SAZ to another).

We further abstract the activity of individual SAZs to proxy measures called energy, density, and flow (EDF). Informally, energy is the sum of motion over zones, density is the distance between zones with motion (how "spread out" activity is around the living space), and flow is the exchange of energy between adjacent zones. The complete definition of EDF is covered in previous publications describing TM (Romero, Pousman, and Mateas 2007). We associate EDF with specific sets of SAZs called regions (such as: kitchen, living, dining, and transit areas). Additionally, we compute "global" EDF over the complete SAZ graph. With EDF values assigned to each region, we have a complete set of proxy measures for human activity in the living space in terms of an alien thought process.

Clustering

The continuous space of values for EDF in different regions is still too complex a model to be expressed in TM. Furthermore, individual EDF values cannot represent long-term patterns in time – the very kind of patterns we would like the system to detect in order to support long-term interaction with the audience! Accordingly, we map the continuous EDF space into a small space of discrete models using an online, soft, *k*-means clustering process. Because the location of cluster centers is updated over time as the cameras observe more activity, they are a function of the entire history of the system and can begin to address long-term patterns.

To support a range from simple to complex behavior, the clustering process operates in several different ways, resulting in different measures which we organize into levels based on their complexity.

Level 1 (L1) bypasses EDF and is based on the activity of a special SAZ associated with the physical space directly in front of TM's display. Using two clusters in a one-dimensional space gives us the effect of an adaptive threshold that moves to separate the natural breaks in the data. Thus the authors do not have to pick fixed thresholds at design time without knowledge of what data the system is likely to see. Clearly, this classification admits descriptions of the L1 state of the form "The special zone is active."

Level 2 (L2) builds explanations of the living space using global EDF. Global energy, density, and flow are independently clustered to produce three "high" or "low" labels. There are $2^3=8$ combinations for these labels – just a few more than we intend to express. We group the results into five cases by merging combinations with high global flow, and treating the other combinations as distinct. We call the resulting case the L2 state. This affords statements like (in the case of high energy and high density) “the system thinks the house is active with all of the activity together” or (high flow) “The system thinks the house is changing states.”

Level 3 (L3) incorporates the per-region EDF and the time of day into a more complex model. Distinct from the other levels, the clustering process in L3 works in a high-dimensional space. Regional EDF contributes fifteen dimensions, and the time of day two more. We use the sine and cosine (with a 24-hour period) of the clock time so that, geometrically, times that are close together during the day are close together in the clustering space. We randomly initialize and iteratively update 32 clusters in this seventeen-dimensional space. The active cluster is called the L3 state. While this only affords statements like (in the case that cluster-17 is active) “the house is in state seventeen,” we have designed the space so that these clusters can find their way to common activities. That is, it is possible that the system could behave in a manner consistent with statements like “The system can tell we are sitting down to watch our favorite television show,” but only because it has a model of what regional EDF looks like during the time of day that the particular show is being watched. However, if no interesting patterns are discovered, it should be easy for the audience to write off the resulting behavior of the system as “just more randomness” as opposed to “acting incorrectly” (though we risk them describing *too much* of the system’s behavior as random).

L1, L2, and L3 are designed to support behavior with different levels of complexity and ease-of-explanation from outside of the system. L1 updates quickly in response to audience provocation in the special region in front of the display. L2 responds to global activity more slowly. Finally, L3 responds to recognized patterns only over very long periods of time (as cluster centers adapt). These different levels of models each correspond to beliefs the system has about its environment arising from an alien method of perception.

Expression

The expression component of TM is significantly more direct than the interpretation component. Here we will describe how the space of L1, L2, and L3 states get mapped to particular images. Figure 3 puts the whole expression component together in a single view.

The L1 state was the simplest model produced by the interpretation process, and here we map it directly to a simple, visually prominent output. When the L1 state reads high, TM goes into “interactive mode” where new images are selected very quickly, causing the display to fade from one composition to the next after only about one second. When the system is in “normal mode”, it selects a new

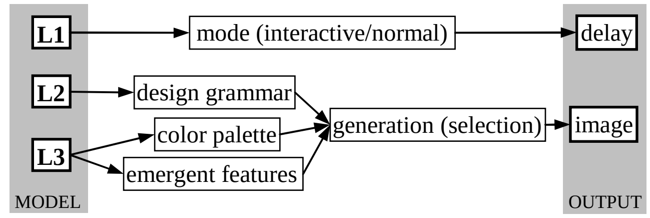


Figure 3: Expressing the state of the system’s beliefs.

composition about once every two minutes. We give the system this behavior so it has a way to say, in its alien language, “Yes, you are engaging me.” Here we have mapped a simple belief to a simple output, trying to avoid leaning on connotations that our alien might not understand. We adopt this “impedance matching” heuristic to (at least attempt to) avoid having the audience become disengaged because the system failed to conform to expectations of complexity, as well as avoiding hiding too much of the system’s potentially interesting interpretation from them.

TM’s L2 puts a basic requirement on the images selected for display. For each of the five L2 states, our mapping dictates that a specific design grammar be used to generate the output image, prescribing a distinct visual style. Figure 4 samples images from the five grammars.

Though we, as authors, have our own aesthetic reasons for mapping certain L2 states to certain grammars, the system is only aware that distinct L2 states are represented with distinct grammars. To attempt to embed any more meaning than our own weak connotation in the mapping of grammars would push on our impedance matching heuristic because the system could be accused of “using words it does not understand.” The alien part of the AP context shines here because our alien is not obligated to be faithful to any particular human interpretation of the shapes produced by each grammar.

The design grammar does not dictate the entire appearance of the final images displayed. The distinct emergent visual properties of an image as well as which color palette is used to embellish its design is controlled directly by the L3 state of the system. In addition to the visual style imposed by the grammar, the L3 state prescribes the “coverage”, “balance”, and “concentration” of images. Coverage is the property of how much the foreground covers the background. Balance describes whether the foreground detail is left-heavy, right-heavy or balanced. Concentration, similarly, describes whether the foreground detail was center-column-heavy, side-columns-heavy, or not distinctly one way or the other. Each of the 32 distinct possibilities for the L3 state is mapped to the presence or absence of three different visual features (and one of four color palette families).

The result of the expression process so far is a grammar name, a set of visual features and a color palette (along with the update rate from L1). As implemented, this is only one step away from the outputs dictated by the agent model. We turn these requirements into a concrete image using the generation component of TM. This image is then output on the display where the audience is free to “decode” it. In the next section we will look at the generation component in more detail.

The expression component, as a whole, translates a simple set of beliefs into an alien (but presumably partially understandable) language. Where possible, we designed the expression component so that association between human-level activities and system’s behavior is fuzzy enough for randomness to wash away any “mistakes.” Recall that we aim to support and engage the audience’s meaning-making, not to have the system declare its own interpretation of the environment to be absolute truth.

Generation

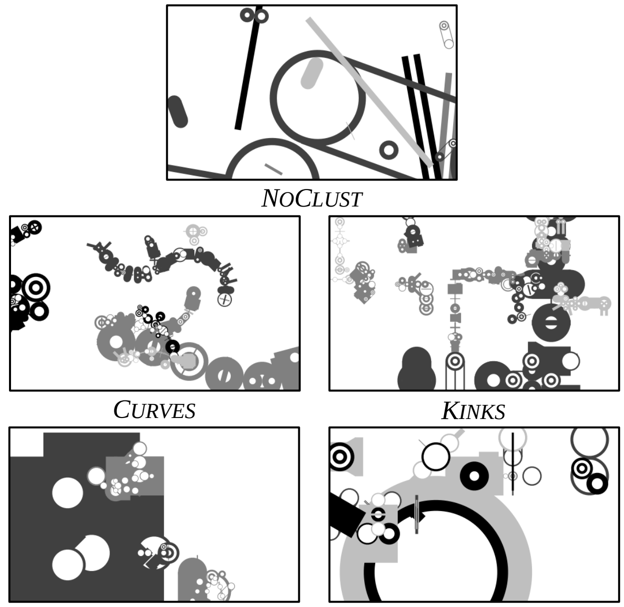
For the system to support any meaning-making at all, TM must display images that are at least engaging enough to spark investigation, and the properties controlled by the L2 and L3 states must be distinct enough that the audience notices the distinctions and can begin interpreting them. These constraints, intersected with the authors’ aesthetic motivation to match the style of our human-designed prototype images, set high expectations for the generation component.

In an idealized world, we would have a simple method to directly synthesize an image each time the system wanted to express its many states. After exploring other avenues, we settled on a powerful and unintentionally general generate-and-test method for achieving our combined goals. The “generate” part uses design grammars to over-produce a space of images that includes those we are looking for as well as many others. The “test” part uses basic image processing to assign labels to images that we can use to filter out only appropriate images. Taken together, these independently-tweakable parts of TM’s generative component comprise a high-performance, parameterized image synthesis process.

“Generate”

We generate images using context-free design grammars. Specifically, we use Chris Coyne’s open-source CFDG package from http://contextfreeart.org/. Informally, design grammars in CFDG are sets of simple rules describing how to draw shapes in terms of other shapes. Rules may be written in terms of primitive (terminal) shapes such as circles, squares, and triangles, in terms of other rules, or even in terms of themselves in the case of recursive rules. Furthermore, several rules may share the same name, indicating that there are several ways to draw the named shape. This practice yields a non-deterministic grammar. This non-determinism, coupled with exploring a huge space of random seed values, is what gives rise the immense space of images that we select from in the generation component.

The main design grammars are built using a stack of grammars, comprising a shared library of TM-specific



NoClust *Curves* *Kinks* *OuterClust* *InnerClust*

Figure 4: Example images from the five design grammars.

shapes, to mitigate complexity and enforce a common visual motif. Each contains rules that describe the overall placement of grammar-specific high-level shapes, as well as the definition of those high-level shapes in terms of the rules from a grammar called Any that describes the smaller shapes common to all of our main grammars. For brevity, we will admit only a brief description of each main grammar (samples shown in Figure 4).

The simplest grammar, NoClust (used for high-flow L2), haphazardly scatters Any shapes, making use of minute angular offsets and gross scaling to give a disheveled look.

The condensed-looking grammars (used for high-density L2), Curves and Kinks, are populated by worm-like shapes composed of long chains of Any shapes that flip orientation and heavy-vs.-lightness along their length according to an improvised, first-order Markov model.

The gaseous-looking (used for low-density L2) grammars, OuterClust and InnerClust, are populated by composite clusters. In OuterClust, clusters are made by growing a seed shape and surrounding it with smaller, similar shapes of the same color, creating a bubbling silhouette of familiar primitives. Alternatively, in InnerClust, clusters are created by nesting shapes of differing color at increasingly smaller scales inside of an outer shape.

Each grammar describes images that are quite distinct from other grammars; however, within each grammar there is still an effectively infinite space of variation. The arrangement of shapes in a final image is the result of sampling from the generative space of a design grammar using a specific seed value for the internal, randomized rule selection processes.

The rendering system for our design grammars is capable of producing high-quality, full-screen images in a vector (shape-based) image format. While we use this format for display, we will see that we will have to generate raster (pixel-based) versions of the compositions to support automated analysis of their content in the "test" process.

"Test"

Recall that the expression component mapped the L3 state to distinct visual properties of images, not just to the name of the grammar that generated it. In order for the system to have a better idea of what the compositions it generates "look like", we pass low-resolution, raster images to an image analysis program. This program looks only at very basic properties of a foreground-background map (independent of intensity). This process results in numerical assignments for the three visual features mentioned in the expression component (coverage, balance, and concentration). These values are in a continuous space, so we chose a working threshold via inspection and used it to assign categorical labels for each feature.

A pixel-level analysis of the images is important because many visual properties are not obvious from a shape-level description of an image. For instance, an image with a single large shape obscuring many small ones appears to be a very simple composition at the pixel level, however the shape-level description would suggest a complex result. Alternatively, if all of the shapes in an image happen to cluster together on the left half of the image, the viewer may perceive a distinct imbalance that is, again, not obvious at the shape-level.

Time-Space Tradeoff

When given enough time, our generate-and-test process can produce an image suitable for expressing any state the system is in. However, by design, TM is a soft real-time system that depends on meeting deadlines to support live human interaction. Because of this constraint, we chose to run the generate-and-test process off-line and save the results.

Because the soft real-time parts of the system only require that an image be produced with a set of requirements (coming from a finite space), we can pre-compute a large set of images for use in each state the system could be in. In practice, this meant sampling about 26,000 images from each of the five main grammars. After analysis, the bulky raster version of each image was thrown away. We saved a compressed version of the vector source for images along with the result of analysis in a database and used this database in a read-only manner for live installations. In this way, the deployed system need not include the ability to sample from a grammar or analyze images, greatly simplifying the software aspect of our live installation.

Clearly, in terms of the visual output of the system, whether the generation process occurred in the home or in the studio before installation is not important (the process does not learn from experience). In either case, suitable images are produced. In terms of the interactive nature of the output, our choice was critical, as, for a given set of requirements on an image, it may have taken hours of search to find a suitable image in the generative space of a design grammar.

Finally, to make our "lookup table" more effective, we restricted the generate-and-test process to grayscale images and left the final application of color palettes (a computationally simple process) to the on-line part of the system. In this way, our database-driven image synthesis process could use any of over 2,000,000 unique, highly-detailed compositions when synthesizing a concrete image from abstract requirements (with only about 50,000 images shown over the lifetime of a TM installation, most going unobserved). Figure 5 shows a final, colored composition.

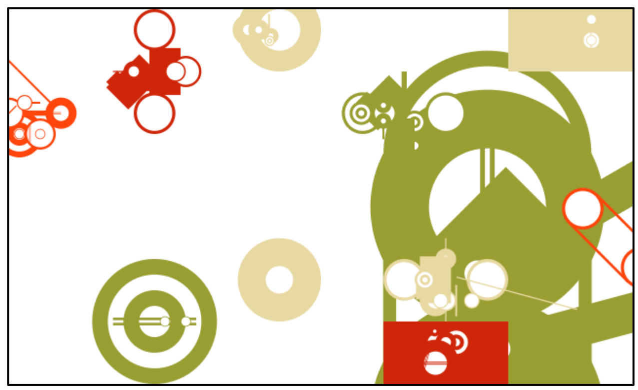


Figure 5: A large composition from the OUTERCLUST design grammar, colored for final display.

Discussion

We have seen the detailed design of TM, from input video streaming from cameras in a living space to a smoothly fading, detailed image sequence at its output. Now we can move to a higher level of discussion and look at our original goals, implications for APs, and creative intelligent systems in general.

Goals

First, we set out to produce a real, working system. The complete TM has been installed and evaluated in homes. In these installations hundreds of images were printed and tens of them annotated with the audience member's thoughts. TM has also been on (interactive) display at the Beall Center for Art and Technology at University of California, Irvine where it was warmly received by artists as well as the general public.

Towards our goal of understanding the design flexibility afforded by the non-anthropomorphic focus of APs, we found the "alien" metaphor quite powerful in the design evolution of our system. After much debate about how to use connotation in the expression component of TM, we realized that an alien need not understand the connotations we were trying to embed. This allowed us to focus on simply illustrating distinctions with distinct outputs, instead of leaning on a model of human emotional response (or the authors' approximation to that model). The focus on our alien's body and its visual input and output organs (cameras and display) allowed us to sidestep hard problems in natural language processing, opting instead for a greatly simplified alien language. Finally, designing TM as an AP allowed us to focus on the intelligent agent model, guiding us to view the system as choosing actions in response to its model of the environment (affording more structure and meaning than viewing it as a plain generative system would provide).

In the realm of human-computer interaction our goal was to understand how TM engages human meaning-making in a long-term interaction. In order to evaluate this question, TM was installed in three homes in the Atlanta, Georgia area for a period of six to eight weeks. Each household contained a nuclear family, with both parents and children, along with pets. At the beginning, while participants were confused and a bit confounded by TM's display and behavior, they regarded it as a curious artifact. As time progressed, participants began to notice trends in the images produced, noting "morning" images and "afternoon" images. Participants found many of the images aesthetically pleasing, and used statements like "nice" and "pretty" to describe images that they chose to print. Participants were sensitive to the design grammars describing *NOCLUST* as "busy" and "blocky," CURVES as "like caterpillars" or like "a rosary," (some curves terminate in crosses giving them a rosary like connotation), and OUTERCLUST as "like bubbles." The visual aesthetic was understood and well received by study participants, and many participants formed interesting long-term interpretations of the system's behavior in terms of their own behavior.

Finally, in our generative art goals of ensuring that images are aesthetically valuable and relevant to the situation, as shaped by the authors' vision, we find that TM meets with mixed success. The generative space of images we define for TM is quite large, and easy to sample from. However, getting the images we would like in a specific scenario was difficult. Our choice of context-free design grammars gave us fine, expressive control of prescribed, local structures in the image (such as how a shape's size, rotation, position in relation to other shapes in a chain, or how intricate arrangements of small shapes form a single, larger shape). However, we had no control over emergent, global structures in the images (such as the accidental arrangement of shapes into larger, recognizable structures, the overlapping of independent shapes, or crowding of an image's border). We addressed this lack of global control with the image analysis process, allowing us to regain control over some emergent properties of images (without complicating the design grammars at all). For the simple properties we implemented in our image analysis process, the (offline) generate and test loop was able to find images with the appropriate emergent properties. However, if we were to instruct the analysis process to look for much more specific emergent properties (say, accidental arrangements of shapes that looked like "human faces" or "safety pins", properties that study participants found quite evocative) we might have to exponentially increase the number of images we search before finding a suitable image.

Alien Presence

With an eye toward AP in general, the interpretation and expression components of TM have some reusable parts. TM used online clustering to map the continuous space of EDF into a finite set of easily described models. In other AP systems, we may not be working in a space where EDF makes sense, but in its place would likely be another continuous proxy measure for interesting properties of the environment. Online clustering could, again, be used to produce simple labels, relevant to the entire history of the input parameters, which are readily mapped to expression pathways of similar complexity. Next, TM built its output image by selecting randomly from a pool of images with known-to-be-appropriate visual features. The particular design grammar basis for "generation" and image processing "test" factors, again, may not be appropriate in the context of a different AP. However, the move to use random selection from a set of options that satisfy required constraints is quite general. Systems designed with this pattern may be made to produce quite complex and expressive outputs.

Machine Creativity

With so much AP-focused discussion in this paper, we should return to the discussion of machine creativity and take a creative intelligent systems view of TM before we close.

TM addresses the search for novelty in two ways. First, in the adaptive interpretation, the system looks for *new*, statistically significant, patterns in the data as they arise (though patterns are described in mostly low-dimensional spaces). Next, in the generative component, we support novelty in the system's output by generating such a large space of images that no repetition of images was observed in long-term installations. This may seem like attempting to get through on a technicality (trivial novelty); however, we expect the audience to complain about a lack of conceptual novelty due to a small number of design grammars long before they complain of images repeating.

The search for value is addressed in a similar manner. In combined analysis and selection process, the system does a significant amount of work to find, from the huge space of compositions we consider, only those images that posses certain visual features relevant to the situation at hand. This could be thought of as a notion of value relative to the system's intention to express its current state. While the processes that implement this idea are technically disjoint, the system is always able to find the kind of image it needs, and does this under time constraints. Next, in the *a priori* selection of the space of compositions, the authors embed a static view of their aesthetic sense. The authors' search through the space of interesting design grammars was an intense, human creative process. Nonetheless, the system was able to wield the results of this search in a way that retained the original value found by the authors.

If we were to look into the code defining TM's behavior, we would not find any answers to general questions like "What should a creative system do?" but the system's successful development and installations do indicate that an AP can successfully embody and dramatically amplify the authors' guess at "What might an artistic, alien other, sharing our home do?"

Future Work

The success of a large system like TM bodes well for favorable exploration of related ideas. Here we will cover a few new research directions which directly build off of our experience with TM.

First, a continued investigation of designing in the AP context is warranted. An alien can behave intelligently without having to overcome any number of AI-complete problems. Future APs may forgo TM's camera and display system in place of microphones and speakers, or even other alien senses such as sonar and infrared vision. APs at very different spatial and temporal scales could be explored as well – towards the scale of neighborhoods or just a single desktop, and a few days to a few years.

Next, to improve upon the "over-produce, then filter" synthesis method used in our system, future work should explore analysis of partially produced artifacts as a means for guiding the design grammar sampling process. The generation process in TM searched through the generative space of the design grammars by trying a large number of random seeds sequentially. We can imagine a system that, instead, attempts to search the "most promising" areas of the space first, detecting the emergence of global patterns before they have fully formed, eschewing otherwise fruitless computation.

Finally, the design of TM addressed key issues in the discussion of machine creativity in a rather accidental manner. Instead of addressing some necessary conditions for creativity ("treating the symptoms"), future work could adopt a particular theory that provides sufficient conditions for creativity and directly implement these. This is no trivial task; however, being able to choose the particular domain for an AP designed with this intent does make the endeavor more approachable.

Conclusion

We have presented the detailed design of *Tableau Machine*, the first alien artist in the home. We showed how TM builds models of its environment to varying levels of complexity and expresses these models in a long-term interaction with a human audience using the relevant output of a prolific generative art system.

We have also tackled the problem of constrained image generation in a concrete setting, addressing both authorial intent as well as soft real-time constraints. The solution in TM uses an internal analysis process to form a primitive understanding of what the compositions it displays look like, and uses this understanding to make decisions about its output.

We hope our experience with TM stimulates further exploration of the space of generative art systems that interact with their environment over extended periods of time.

Acknowledgements

This research has been supported in part by a grant from the National Science Foundation, IIS-0534355.

References

Buchanan, B. 2001. Creativity at the Meta-Level: AAAI 2000 Presidential Address. *AI Magazine*, 22(3), Fall 2001.

Cohen, H., 1995. The Further Exploits of AARON, Painter. In *Constructions on the Mind: Artificial Intelligence and the Humanities, Special Edition of Stanford Humanities Review* 4: 141-160.

Cohen, H. 1999. Colouring Without Seeing: A Problem in Machine Creativitiy. Retrieved January 25, 2008 from http://www.kurzweilcyberart.com/aaron/pdf/colouringwithoutseeing.pdf.

Machado, P., and Cardoso, A. 1997. Model Proposal for a Constructed Artist. In Procs. of the First World Multiconference on Systemics, Cybernetics and Informatics, Vol 2: 521-528.

Mateas, M. 2001. Expressive AI: A Hybrid Art and Science Practice, in Leonardo: Journal of the International Society of Arts, Science, and Technology. 34(2). 147-153.

Pease, A., Winterstein, D., and Colton, S. 2001. Evaluating Machine Creativity. In First Workshop on Creative Systems, International Conference of Case-Based Reasoning.

Romero, M., and Mateas, M. 2005. A Preliminary Investigation of Alien Presence. In Procs. of Human-Computer Interaction International (HCII 2005).

Romero, M., Pousman, Z., and Mateas, M. 2007. Alien Presence in the home: The Design of Tableau Machine. In Journal of Personal and Ubiquitous Computing. 12(5).