

DETC2009-87702

## SUPPORT FOR THE USE OF HIERARCHICAL TEMPORAL MEMORY SYSTEMS IN AUTOMATED DESIGN EVALUATION: A FIRST EXPERIMENT

**Josh Hartung**

Department of Mechanical Engineering  
University of Idaho  
Moscow, Idaho 83844, U.S.A.

**Jay McCormack\***

Department of Mechanical Engineering  
University of Idaho  
Moscow, Idaho 83844, U.S.A.

**Frank Jacobus**

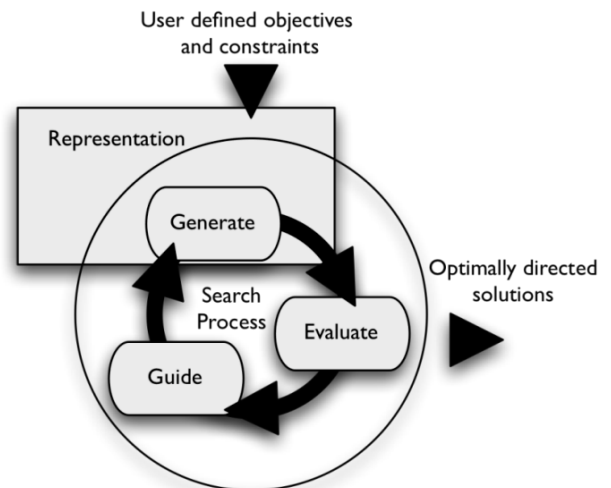
Department of Art and Architecture  
University of Idaho  
Moscow, Idaho 83844, U.S.A.

### ABSTRACT

Computational synthesis tools that automatically generate solutions to design problems are not widely used in industry despite many years of research. This deficiency can be attributed to the lack of value that these tools provide for the user in terms of time saved in design or quality improvements in the design. In order to provide sufficient quality of solution, it is proposed that more human-like evaluation of solution quality is needed including qualitative concepts, the ability to allow for anecdotal input, and general inclusion of ambiguous information. A hierarchical temporal memory system (HTM) is proposed as a viable approach for capturing design quality from exemplars and subsequently recognizing the presence of that quality in other designs. This paper includes a first experiment in using HTMs for learning and recognizing quality in the form of the visual style characteristics of Hepplewhite, Stickley, and Greene & Greene chair backs. Results show that HTMs develop a similar storage of quality to humans and are therefore a promising option for capturing and recognizing multi-modal quality information in future design automation projects.

### 1. INTRODUCTION

The engineering design process is the most expensive and time-consuming phase of product development. Designers must first understand the needs and constraints of the customer and develop an adequate representation of that information.



**Figure 1.** A graphical model of computational synthesis as described by Campbell, et al. [1]

Solutions can then be generated in an attempt to fulfill those needs. Typically multiple iterations are required to ensure customer satisfaction. The cyclic process of design and redesign is both time intensive and highly dependent on the skill of the designers. Computational design techniques have attempted to mitigate these problems by automating part or all of the process.

The computational synthesis process can be represented as a cycle of generation, evaluation, and guidance [1] (Figure 1), working within a representation of the problem and design space. The generation process first forms potential solutions to all or part of the design problem. The quality of the solutions is then determined by an evaluation routine, which guides further refinement of the solutions. Most computational efforts have been applied to creating and adapting generation methods, such as stochastic optimization, genetic algorithms, expert systems, etc. By narrowing the scope of a problem, computational techniques have been successful in automating some elements of the design process, but these tools are less effective when working in broad domains or with ambiguous data.

Despite research advances in computational synthesis, design automation tools are not widely used in practice because they do not sufficiently reduce design time - degradation in quality typically accompanies automation. For example, automated placement and routing technology used to design integrated circuits has been adopted because of speed improvements on the order of 10 to 100 thousand times[2]. Most would acknowledge that designers working by hand could still “do better” or at least the human perception of the results would “look better”, but the time saved is so great that the automation is enabling in the creation of these products. When other computational tools offer speed improvements that are only a factor of 2 to 10, the solutions must be perceived as near human quality. To improve the computational design process, new methods are needed that can address the persistent problem encountered by automated synthesis tools: computational design systems are not effective at evaluating comprehensive design quality. The ineffectiveness can be attributed to challenges in three areas:

- Representing difficult to parameterize and quantify aspects of quality
- Representing and reasoning about ambiguous data
- The preference of users to express quality through anecdotal data

A human designer easily operates using ambiguous data and has an innate sense of quality during design. Quality can be evaluated from the outset, as the human considers disparate specifications or user objectives, and is continually evaluated throughout the process. Without the ability to design with an understanding of quality, computational techniques often produce solutions that designers feel are limited. If a customer desires a printed circuit board that is both functional and consistent with their personal layout style, they will generally reference examples of past designs. A human designer will maintain their stylistic properties as global considerations that affect both generation and evaluation. Computationally, the concept of layout style has no generalized representation, and therefore cannot be considered when generating solutions.

In this experiment, we utilize a computational model known as hierarchical temporal memory[3] (HTM) in order to store and interact with a model of quality that is analogous to that of humans. HTMs are built on a model of human neurological function and have shown great promise in areas of ambiguous pattern recognition in vision and speech applications. Most significantly, they have the ability to store data on a series of training patterns and subsequently identify incomplete or out of sequence pattern data that they have not seen before. This allows the system to learn the characteristics of a series of anecdotal data and then evaluate new data based on its similarity to the exemplars. In a design context this is a first step toward a technique for computational evaluation of quality by analogy, where solutions are judged based on their similarities to known, preferred solutions. The ability to find meaningful connections between groups of objects while allowing for differences and computing with ambiguity means that new solutions can be dramatically different and even novel while retaining some core familiarity that is preferred by the user.

As a first test of the capability of HTM systems in representing and recognizing quality, an experiment was performed which compares HTM performance in characterizing and identifying chair back style (Figure 2) to the performance of humans. Quality is defined as “the standard of something as measured against other things of a similar kind”[4].



**Figure 2.** A chair back

Therefore, computational implementations of quality must consist of a representation of a standard and the measurement against that standard. The experiment consists of a standard defined by exemplary chair back styles where the characteristics of the styles are learned by the HTM. The measurement is performed by comparing proposed designs to the standard. After developing an understanding of three chair back styles, nine novel designs are presented and the machine is asked to categorize them. Finally, a sample of architectural students with no previous exposure to these designers is asked to do the same task and make some qualitative statements justifying their choices. A high degree of consistency is found between the machine and human results. Further interpretation of the category data from the machine and the qualitative statements from the human subjects supports the validity of a computational representation of style that is consistent with a human one.

This experiment provides support that HTMs can be effective tools for developing a model of quality. Human designers must first learn to recognize quality before they can evaluate it, and they must learn to evaluate it before they can create it.

Once computational tools can recognize quality, evaluation can be enhanced. A computational model of quality that is consistent with that of a human is a powerful first step toward general automated design tools.

## 2. BACKGROUND

The established measure of quality in design synthesis comes in the form of cost functions, objective functions[5], and utility functions[6], which are mathematical representations of the goodness of a solution or a user's satisfaction with a solution. These functions operate by encapsulating and aggregating one or more design characteristic into a single measure. Difficulties with this approach arise when there are multiple objectives to represent. Weighting is typically applied to multiple objectives to align the relative importance of each variable with respect to one another. Unfortunately, discovering a set of weights that are uniformly applicable to a design problem is difficult. This can be further complicated when the weights are analogous to preference because multiple users may differ in terms of their definition of quality for a product. In a commercial system, these approaches require the developer or user to adjust weight values, which is impractical and unintuitive for the end user that simply wants good results.

Another perspective on quality and design tradeoffs is the Pareto surface, which contains solutions that cannot be improved in one aspect without worsening another. This is notable with respect to quality in that it illustrates the uniform quantitative characterization that objective function-like approaches will apply to vastly different solutions. It is unlikely that a user will be equally as satisfied with all solutions along the Pareto surface. There is typically outside information that distinguishes good from bad solutions that was either difficult to weight, express quantitatively, or even to verbalize.

Other methods such as conjoint analysis[7] have been used to extract preferences by asking users to make trade-offs with respect to design features or parameters. Trade-offs are gathered through surveys in which users rank or rate carefully chosen design options that are selected to elicit their preferences, and the results can then be codified in quantitative measures. With this approach the designer must construct the set of design options to be used in the preference surveys for deriving quality versus directly using examples of user quality.

Other relevant research has been performed to enable computational systems to reason using ambiguity for decision-making, robotic control, or design reuse. Concepts such as fuzzy logic[8], case-based reasoning[9], and rule-based systems such as shape grammars[10] use the capability to match a source example to target problems to reuse knowledge (case based reasoning) or determine the applicability of a rule. Similarly fuzzy logic outlines a mapping between inputs and responses in which a range of inputs can map to a single response, thereby eliminating the necessity for an input to be precise. In these systems, tremendous work is required to

define the rules by which data can be transformed to match exemplars of similar form (i.e. geometry, graphs, etc.) or define the appropriate ranges of response. Even though no single approach yet provides the desired capabilities in representing and recognizing design quality, existing technologies including HTMs illustrate the potential to create robust computational measures of quality.

### 2.1 HTM SYSTEMS

Hierarchical temporal memory (HTM) developed by Numenta, Inc. [11] is a recent form of artificial intelligence that excels at ambiguous pattern recognition. Based on an emerging theory of human neurological function, they can solve ambiguous and multi-modal problems which traditional computing systems find difficult or impossible. Promising results have been observed in machine vision, voice recognition, and objective character recognition[12].

HTMs are not programmed in a traditional sense; rather they are trained on input data. They have two modes: training and inference. During training, learning nodes arranged in a

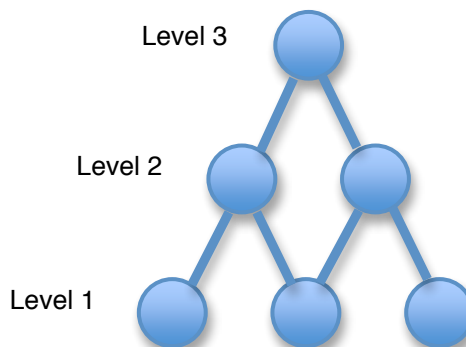


Figure 3. An example hierarchy of learning nodes

hierarchy (Figure 3) identify and store patterns in space, and then in time. Figure 4 details the flow of data within an individual node. Vectorized information is passed first into a spatial pooler where it is evaluated for similarity to other vectors that have already been seen. If the vector is spatially

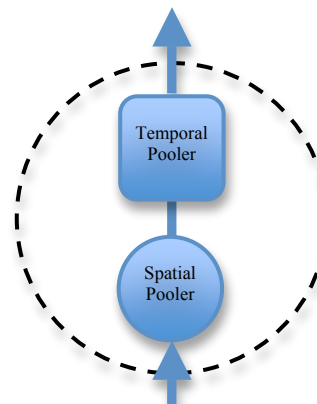


Figure 4. The flow of data within a learning node

the same or similar (within some defined range) to a vector that has already been seen, it is identified as a coincidence of the exemplar vector. If the vector is beyond the defined range, it is recorded as a separate coincidence. The coincidence information for each vector is then passed to a temporal pooler.

The temporal pooler identifies patterns in time. As new vectors are received, coincidences come from the spatial pooler in a sequential manner. These sequences are stored and evaluated for coincidence in a similar manner to the spatial pooler. Finally, the temporal coincidences are output from the node. This information is then propagated up the hierarchy (Figure 3), each level feeding its parent the information from several nodes. The result at the highest level is an invariant understanding of the problem data, which may be used for inference.

During inference the HTM is presented novel data from the same category as the training data. This data may be highly ambiguous. For instance, if an HTM was trained on photos of full cows, then novel data might be a picture of a cow, but partially obscured behind a barn. This is a completely different data set than was initially presented: large amounts of data are missing. Perhaps the head of the cow, the hind legs, or the midsection are obscured. Each of these circumstances presents a problem to traditional pattern recognition systems. The HTM however, still sees a significant portion of the patterns it identified during training. As this data propagates up the hierarchy, incomplete patterns are filled in by probabilistic analysis. The HTM returns the result: a cow. Just as a human would do, the system assumed that because it saw the front of the cow, then the rear of the cow must be attached and standing behind the barn.

In this example, the HTM stored the underlying patterns that represent all cows in all situations from a limited set of data containing a few cows in a few situations. Rather than storing specific information about the form of the animal, the HTM stores the hierarchy of coincidences over space and time that are common to the observed cows. In this way, the HTM stores underlying patterns and uses them to interpret vague or incomplete data in a similar manner to a human.

## 2.2 Neural & Bayesian Networks

Many of the concepts employed by HTMs are not new or unique to them. The novelty and potential of the system is a result of the combined use of existing concepts[13]. Qualities of neural and Bayesian networks in particular are apparent from the description of the system's structure and behavior. A Bayesian network[14] is a graphical representation of probability through which the dependencies of a collection of variables is modeled. The network of variable relationships can then be used to answer probabilistic questions about the variables.

A neural network[15] is a collection of simple processing nodes that together can exhibit complex behavior. Neural networks are primarily used in machine learning applications.

Through a process of training, neural networks can learn to recognize complex patterns of data in order to produce a response.

In the HTM, existing methods are leveraged where they are useful. The distinguishing feature that they offer is the recognition and storage of spatially consistent data over time in a hierarchy. This process allows the HTM to capture an invariant model of the input data, which allows computing with high levels of ambiguity.

## 3. EXPERIMENT

The experiment attempted to use an HTM system to store and recognize the styles of chair backs made by three different designers. Because the HTM is based on human neurological function, it was expected that the storage of chair back style information would be analogous to that of a human. To verify this, human subjects were asked to develop a qualitative representation of style for the chair backs using image data that was identical to that which the HTM was trained on. By correlating qualitative statements made by the human subjects with the results obtained from the HTM, we found the system's recognition of quality to be analogous to that of the human. This provides validation for further use of the system in design automation applications.

### 3.1 Chair back Data

Chair backs were taken from classic examples of the work of George Hepplewhite, Gustav Stickley, and Charles & Henry Greene. Six designs from each category were selected for the experiment (Figure 5). Three of the six designs were grouped together because of their stylistic consistency. Each design in the consistent set shared key stylistic features with the others in that set. The inconsistent set of chair backs consisted of one canonical design and two others that combined elements of different styles per category. For example, consider the Greene chair backs from Figure 5. The chair backs on the left are a consistent sample of the designer's style featuring a strong central element, curved or stepped shoulders, and an ornate center section. The chair backs to the right are a deviation from the canonical style. Specifically, Greene 4 is an outlier because of its ornate center section that is similar to the Hepplewhite styles. Greene 5 is canonical because it features the rounded shoulders and strong central member of the training set. Greene 6 is also an outlier with the strong rectilinear construction of Stickley.

Data was created by hand in the form of black and white silhouette images of a high resolution. Because of processing limitations, the images were scaled to 128x128 pixels using a widely available graphic conversion program. The grayscale dithering of the edges was maintained in order to ensure consistency between the computer and human data, which was not fully black and white either.

### 3.2 HTM Trial

Two networks were used which were based on the Numenta Vision Framework, a version of HTM specializing in image recognition. Both networks were trained on three images from

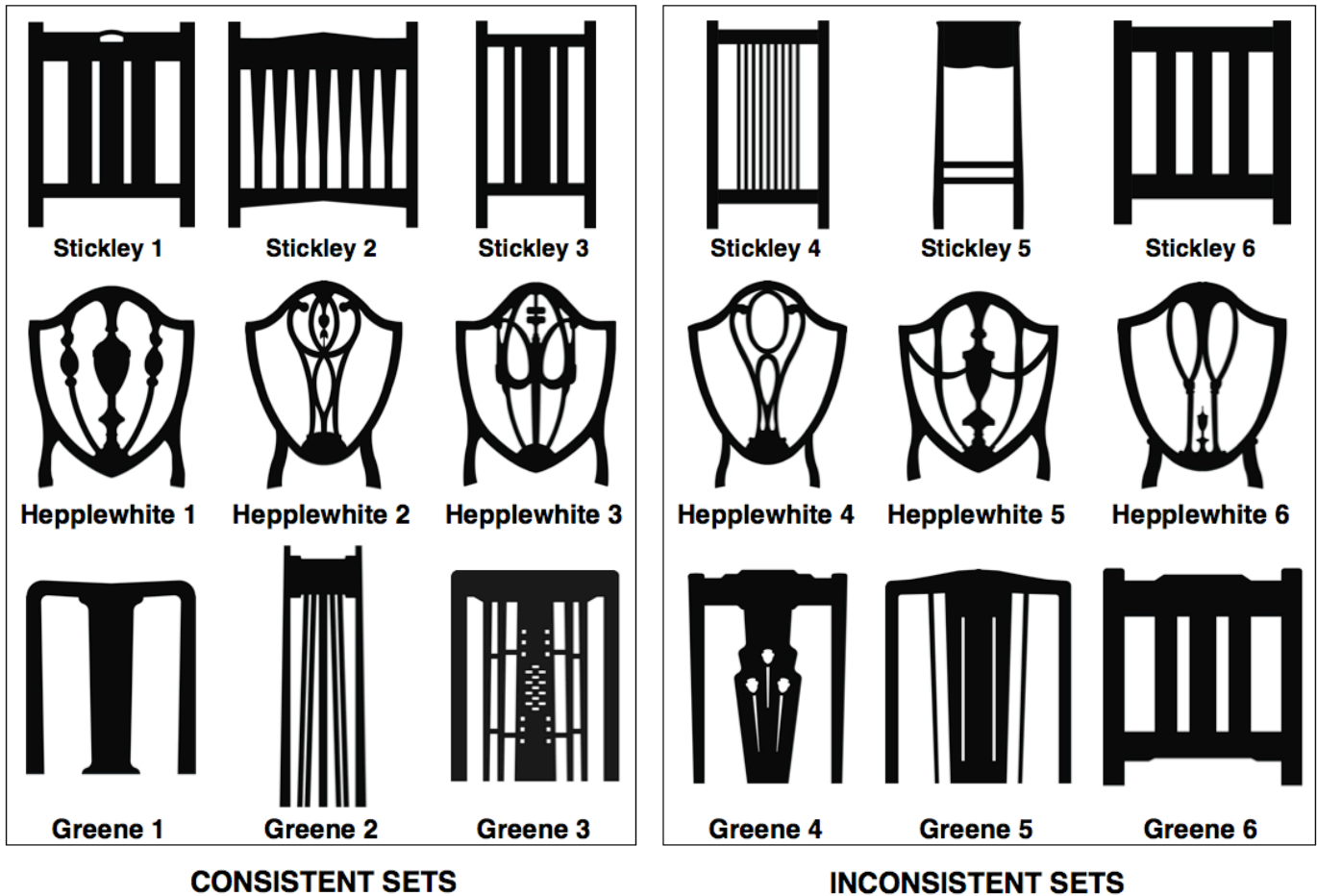


Figure 5. Chair back data

each designer. One network was trained on the stylistically consistent sets seen in Figure 5; the other network was trained on the inconsistent sets.

The system inverted the images and placed them on a black background, eliminating the white background present in the originals. This was done in-situ by the Vision Framework's onboard image conversion features. Network parameters were then tuned by hand to ensure spatial learning was storing unique coincidences and there was temporal coherence between the spatial patterns.

When training was complete, the system attempted to classify all the images of both the training and testing sets (Figure 5). It returned a belief distribution for each image that indicated its first best guess, followed by a second and a third. Figure 6 shows the results of inference for Hepplewhite 4 from the Hepplewhite set. The bar graph at the top right displays the system's belief distribution that the particular image belongs in each category. For reference purposes, an archetypal image of each category appears at the bottom of each bar. In this case, the system has the highest confidence that the sample belongs in the Hepplewhite category, followed by Stickley, and finally a low confidence that the sample is a Greene & Greene.

### 3.3 Human Survey

A human survey (sample found in the appendix) was conducted on condition of anonymity with selected architecture students at the University of Idaho. Architecture students were chosen for their formal training in identification of the stylistic characteristics of design. In order to limit the information external to the survey that was available, care was taken to ensure that the subjects had no prior familiarity with the three designers or their chair back styles. Because the computer has no external knowledge of chair backs, it was necessary to ensure that the subjects did not either.

The subjects were given one of two surveys. The first presented them with training data identical to the consistent set (Figure 5) that the HTM was trained on. The other contained the inconsistent set (Figure 5). They were asked to make at least three qualitative statements characterizing the style of each of the categories. An example was given using a cow, where the style of the cow was described with statements such as "spotted – greater area of light than dark" and "small legs in proportion to large body".

After evaluating the training data, students were presented with the testing images from the HTM trials. They were asked

to provide a 1st and 2nd guess at the category to which the image belonged, a percentage of certainty for each guess, and justifications for their guesses using the descriptive statements made in the training phase.

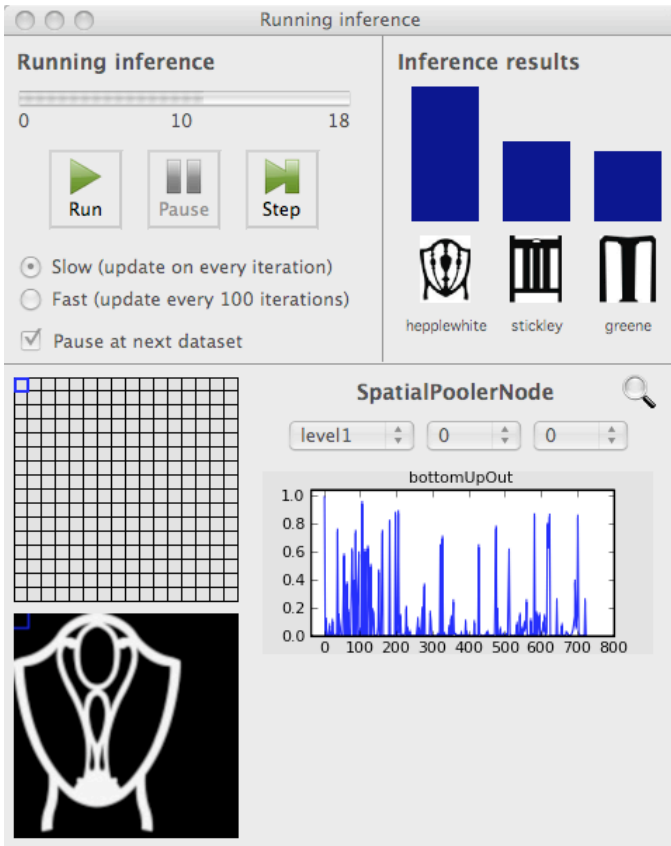


Figure 6. The results of inference on Hepplewhite 4

#### 4. RESULTS

The results from training and testing the HTM are presented in section 4.1 followed by a description of the survey responses from human participants in section 4.2. In each section the results from inconsistent training are presented first followed by the consistent training results.

##### 4.1 HTM

When it was trained on the inconsistent set, the HTM struggled to place images in their correct categories. It correctly placed 6 of 9 testing images in their correct categories and displayed a high degree of confusion for most choices. It also displayed highly non-logical second choices for most cases, such as a high secondary certainty of Stickley for a Hepplewhite chair back. Stickley, with its highly rectilinear design should have a low certainty when categorizing the highly curvilinear design of the Hepplewhite.

When it was trained on the consistent set, the HTM correctly placed 7 of 9 images in their correct category with a high degree of certainty. Greene 6 (Figure 7) was placed in the

Stickley category with a high degree of confidence. Greene 4 (Figure 8) was placed in the Hepplewhite category. In this case the system displays a high degree of confusion, where all categories appear with nearly equal confidence. This indicates Greene 4 shares stylistic features with all three of the categories.

In general, when a testing instance shared many stylistic elements with the training data for one of the category groups the computer had a high degree of certainty that the sample belonged in that category. When selecting the training data, care was taken to provide the computer with at least one canonical sample that fulfilled this circumstance. These images were Stickley 6, Greene 5, and Hepplewhites 4 and 5 (Figure 5). The HTM correctly identified every one of these images with a high degree of certainty.

During inference, the HTM displayed a minimum level of certainty for every category across all samples. There were no chair backs for which the system displayed a very low or zero certainty for any of the categories.

##### 4.2 Human

When provided with the inconsistent training set, humans mis-categorized chair backs and had a high degree of uncertainty about their choices. In this case, the humans were on average less accurate than chance, indicating misdirected storage of invariant information about each chair back category.

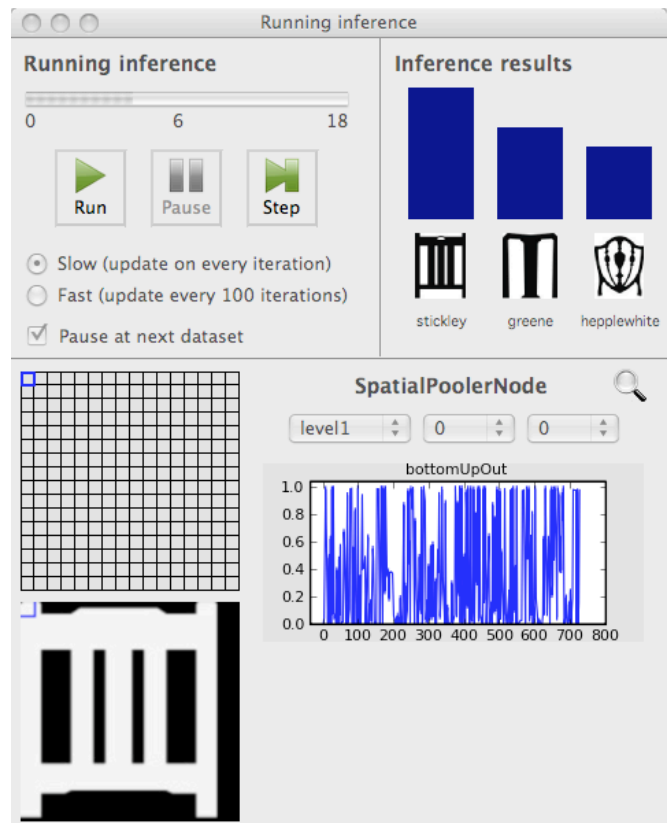


Figure 7. Greene 6

When they were provided with the consistent training set, humans were as accurate as the HTM, correctly placing 7 of the 9 chair backs with a high degree of certainty. As with the HTM, every human placed Greene 6 (Figure 7) in the Stickley category. All but one human placed Stickley 5 (Figure 9) in the Greene category.

Of the five surveys returned for the consistent training group, one subject incorrectly placed samples other than Greene 6 or Stickley 5. The subject placed Stickley 4 in the Greene category.

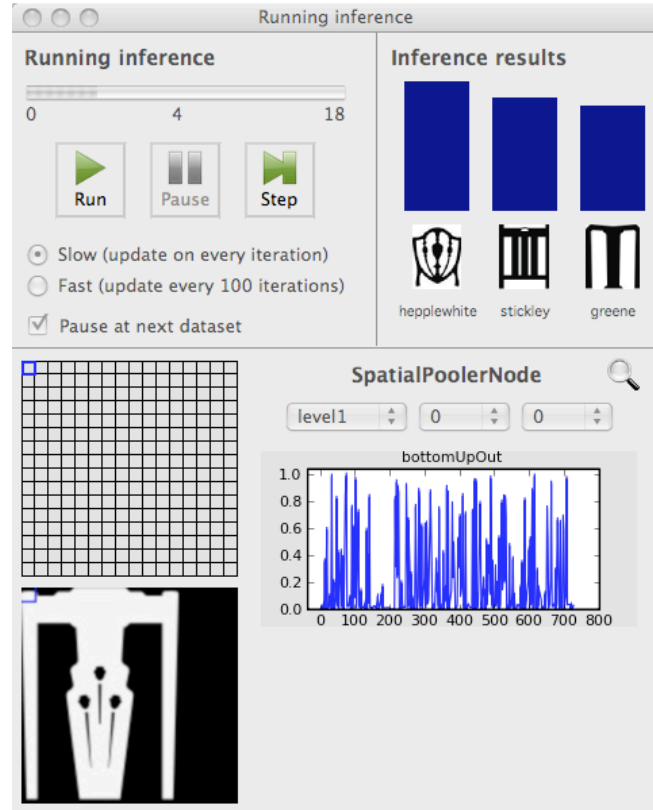


Figure 8. Greene 4

## 5. DISCUSSION

To validate a computational model of quality that is consistent with the human model, the human model must be represented in a way that lends itself to comparison with the computational model. In the survey, subjects were asked to define a style using descriptive statements. By identifying key stylistic features common to all the chair backs within a category, the subjects were forced to display the invariant knowledge that they recognized about each designer’s style. Next they were asked to categorize testing data using two guesses. The survey explicitly instructed the students to justify their guesses “by referencing [their] descriptive statements for each group” (appendix). The goal was to simulate the inference process of the HTM by using only category information stored during training.

The HTMs followed an identical process of training and inference. They initially stored invariant information regarding the style of each chair back category until a sufficient invariant representation was reached. This is analogous to the humans making descriptive statements about each style except that the HTMs display their invariant understanding of the chair back styles through belief distribution for each sample during inference. Analyzing HTM and human inference decisions, it becomes clear that the two models are highly similar, thus providing a firm basis for further experimentation with HTMs in design evaluation and synthesis.

### 5.1 Training with Inconsistent Data

Pattern storage in both humans and the HTM was performed from a set of data that is exemplary of a particular style. By identifying key elements of a style in an invariant form, similar elements in novel data are identified and matched to the training patterns. The result is ambiguous pattern recognition. The robustness of the process depends on the quality and consistency of the training set. In a consistent training set key stylistic elements are shared by all members. In a non-consistent set, stylistic elements vary between members so that they may not be stored in an invariant form that is common to all members. This is how humans differentiate between dissimilar objects. For example, the stylistic elements that are shared between an orange, a coffee pot, and a bicycle are few, thus they form an inconsistent training set. A human would have difficulty identifying another member of this set. An orange, an apple, and a watermelon, however share stylistic elements such as round shape, bright color, and sweet taste. Thus they form a consistent training set for the recognition of “fruit”. A human could easily place another object, such as a melon, in this category.

When they were trained on the stylistically inconsistent sets, both HTM and humans had difficulty determining the category in which to place chair back test images. Humans displayed a very low level of agreement between surveys and confidence percentages were low for either choice. The HTM also displayed a high level of confusion between the choices, but only mis-categorized one chair back more than the HTM that was trained on the consistent set and was therefore much more accurate than the humans.

The HTM achieved higher accuracy because it has the ability to access stored invariant information without contextual interference. Humans are required to interpret the data through a complex hierarchy of connected concepts. This interpretation is what allows humans to see the similarities between disparate objects. Because of this hierarchy, a human understands that a pencil, a coffee mug, and a water pipe all share a common, cylindrical shape, for example. But this contextual filtering brings huge amounts of information from outside the problem into it and can be the cause of fixation and information overload. The HTM can only interpret test data in the context of its training. Thus, the scope of the inference it performs can be artificially limited. In the case of the chair backs, the HTM moves to smaller and less repetitive stylistic

elements in order to make a decision, while the human looks for higher-level correlation.

The confusion between both HTM and humans during inference indicates a similarity in the way information is stored and therefore further validates the HTM model of quality as comparable to the human. Without a complex hierarchy of connected patterns to wade through, the HTM can perform inference at a much lower level than the human. This results in robustness to inconsistent training groups.

## 5.2 Training with Consistent Data

Both the HTM and humans achieved a high level of accuracy when they were trained on the consistent data set. While accuracy is an encouraging result of the experiment, it is essential to verify that the justifications for each decision were consistent between the HTM and humans. By examining the justifications for both the correct and incorrect results of inference, we found increased support for that consistency.

**5.2.1 Invariance Across All Data** Despite high certainty percentages for their first choices by both the HTM and humans, the two systems often differed in their second choices. The humans routinely cited low certainty that each chair back belonged to any second category. The HTM displayed about half the certainty for a second category and slightly less than half for the third. The results in Figure 6 are typical of all correct HTM results – the system always assigned a significant belief value to every choice.

A minimum certainty across all choices by the HTM can be attributed to the identification of the object as a chair back. There is a consistent geometry between samples that identifies them all as chair backs. For example, they share a similar scale and they have a contained, high volume shape that forms the body. The HTM identified individual style categories at one level of the hierarchy and an understanding of chair backs in general at a higher level of the hierarchy. Presenting the system with a chair back whose stylistic features did not conform to any of the training categories yielded results that indicated a high level of certainty that the sample was from all categories (i.e. a chair back), rather than none of them. This was the case with Stickley 5 (Figure 9), which was chosen to be in the testing set because it was an outlier in the Stickley style. Accordingly, the machine displayed near uniform belief across the three categories.

**5.2.2 HTM vs. Human Invariance** One human subject placed Stickley 5 correctly, citing the following descriptive statement defining the Stickley style: “2 connection points at bottom outside edges” (appendix). This was the only subject who identified this stylistic property. Significantly, the subject made an almost identical statement characterizing the Hepplewhite style. This was the only statement that was common to both Stickley and Hepplewhite categories in the subject’s survey.

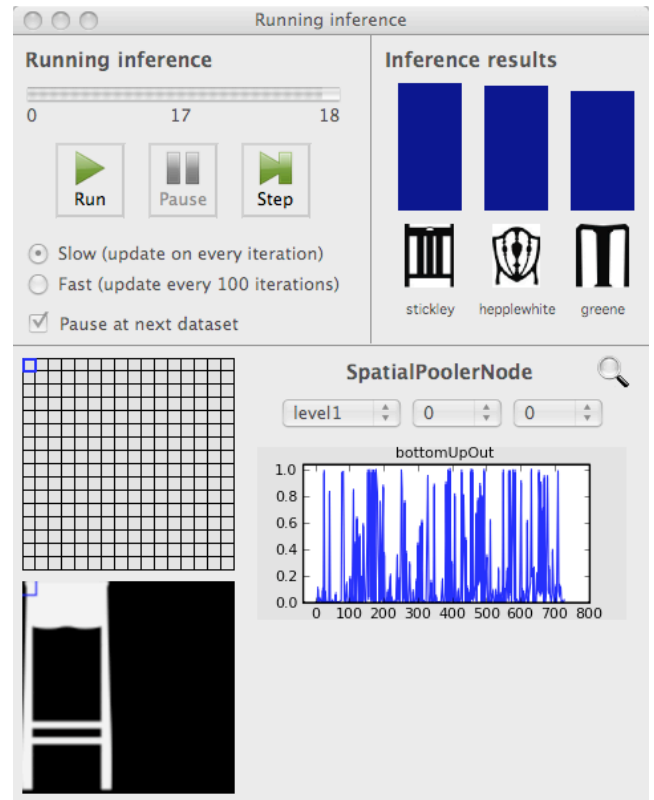


Figure 9. Stickley 5

The HTM placed all the Hepplewhite testing samples into the correct category, but the second choice for Hepplewhite samples was consistently Stickley. Therefore, they both share common stylistic elements. This was consistent with the statement “2 connection points at (the) bottom outside edges” made by the human subject. The Greene training samples do not possess that stylistic feature. It is reasonable to assume that the HTM made a similar justification to the human when making its second choices.

This assumption is further validated by the HTMs response to Stickley 5 (Figure 9). The major defining characteristics of that chair back are a taller aspect, ample empty space contained within two prominent vertical members, horizontal members across the bottom, and two connection points at the bottom outside edges. While every category exhibits at least one of those features, it is clear that Stickley and Greene are the only categories that have only two connection points. Thus the machine placed them with nearly equal certainty in a decision that was analogous to the human subject.

**5.2.3 Attention** During inference, the HTM improperly categorized 2 chair backs. The first was a notable exception: the humans cited reduced certainty in placing Greene 4 (Figure 8) in the Greene category, but the HTM mistakenly believed it was a Hepplewhite. Despite their similarities, the HTM is not a fully realized model of human neurological function. One essential reasoning tool that it currently lacks is an attention mechanism.

When defining the style of Hepplewhite, the subjects overwhelmingly cited the distinctive “shield shape” of the outline and the curvilinear nature of the design. During inference, subjects had no difficulty correctly identifying Hepplewhites whether they had been trained on the consistent or inconsistent set. The two stylistic elements of Hepplewhite are so predominant that they cause fixation, demanding the full attention of the subject. In statistical terms the attention represents weighting of selected patterns, making those patterns more influential when performing inference.

The lack of an attention mechanism in HTMs hampers their performance in situations that humans would find unambiguous. Instead of fixating upon dominant features, the system considers the more minor attributes equally important. In the case of Greene 4 (Figure 8), the system considers the ornate center section and large amount of negative space to be indicative of a Hepplewhite. It does not have the ability to focus on the lack of a “shield-like” outline or curvilinear features to remove the Hepplewhite category from consideration. The future development of attention mechanisms by Numanta, Inc. would greatly increase the robustness of pattern recognition in less ambiguous situations.

**5.2.4 Quality Evaluation** Validating the model of quality stored by the HTM as consistent with the human provides a basis for use of the HTM in evaluative capacities. Computationally, the belief distributions that are derived from inference can be interpreted as an amount to which a certain instance conforms to a broad archetype. Greene 6 (Figure 7), was physically created by designers Greene & Greene, but placed in the Stickley category by the HTM. It was confirmed by the human survey that Greene 6 is more stylistically consistent with the Stickley category. Thus the HTM recognized that the sample conformed more to the Stickley archetype than to the category of its designers Greene & Greene.

In a design context the HTM could be trained on designs that are known to be of high quality by human designers and inference could be performed on novel designs. The results of inference would then display the amount that the novel designs conformed to an archetype known to be of high quality, allowing the HTM to act as an automated quality evaluator using the same process employed in this experiment with chair backs.

## 8. CONCLUSION & FUTURE WORK

This first step toward a general design automation tool lays the groundwork for an expansive investigation into the storage and utilization of quality. By correlating human and computer results of image categorization we were able to validate the HTMs storage of stylistic information as comparable to that of a human. Additionally, the experiment illustrated five points:

- When given inconsistent training data, both HTM and humans have similar difficulty during inference. (Section 6.1)

- When given consistent training data, the both HTM and human inference results are accurate and closely aligned. (Section 6.2)
- The HTM not only learns patterns that exist within a set, it learns patterns that exist between sets. (Section 6.2.1)
- An attention mechanism is essential for the HTM to properly interpret highly distinctive features. (Section 6.2.3)
- The HTM can classify objects only through data that was perceived during training, not through extraneous information. (Sections 6.1 & 6.2.4)

In this initial experiment we chose to utilize a framework for the HTM specializing in image recognition, but the system is in no way limited to visual pattern recognition problems. Almost any data that has temporal consistency can be interpreted by the HTM. Currently other versions of the system are effective at speech recognition, objective character recognition, machine vision, and even statistical trend analysis. In future experiments, we will use the system to analyze the quality of layout of printed circuit boards (PCBs). The physical placement of parts on PCBs is a time consuming process that often lengthens the time to market for new technology and substantially increases the cost. Eventually we hope to utilize the system in a general design synthesis situation, where it can leverage multi-modal analogical data to create new designs that build off of existing ones.

Future work must include the addition of multi-modal input data in order to validate the system for real world design evaluation. In this experiment we limited the scope to simple grayscale images for purposes of validating the model with humans. Real world problems feature complexities that also contribute to their quality such as scale, material, and color, even product function. Future experimentation with the system will go beyond chair backs and for an effective general tool we must strive to go beyond simple visual data.

## 9. REFERENCES

- [1] Cagan, J., Campbell, M., Finger, S., and Tomiyama T., (2005), Framework for computational design synthesis: model and applications. *Journal of Computing and Information Science in Engineering*, Vol. 5, pp. 171-181.
- [2] Applications of VHDL to Circuit Design, Editors – Randolph Harr and Alec Stanculescu, (1991) Pg 128 *Quality/Productivity Design Automation Acceptance Criteria*.
- [3] Hawkins, J. (2004), “On Intelligence” Times Books.
- [4] The Oxford American Dictionary.
- [5] Nash, S. and Sofer, A. (1996), “Linear and Nonlinear Programming” McGraw Hill.

- 
- [6] Keeney, R.L., and Raiffa, H., (1976), *Decisions with Multiple Objectives: Preference and Value Tradeoffs*, Cambridge University Press, New York, NY.
- [7] Green, P. and Rao, V., (1971), "Conjoint Measurement for Quantifying Judgmental Data," *Journal of Marketing Research*, 8(Aug.), pp. 355-363.
- [8] Novák, V., Perfilieva, I., and Močkoř, J., (1999), *Mathematical principles of fuzzy logic*, Dodrecht: Kluwer Academic.
- [9] Leake, David., (1996), "CBR in Context: The Present and Future", In *Case-Based Reasoning: Experiences, Lessons, and Future Directions*. AAAI Press/MIT Press, pp. 1-30.
- [10] Stiny, G. (1980), "Introduction to Shape and Shape Grammars" *Environment and Planning B: Planning and Design*, Vol. 7, pp. 343-351.
- [11] "Hierarchical Temporal Memory – Concepts, Theory, and Terminology", Jeff Hawkins and Dileep George, Numenta, Inc., cited Feb. 27, 2009  
[http://www.numenta.com/Numenta\\_HTM\\_Concepts.pdf](http://www.numenta.com/Numenta_HTM_Concepts.pdf).
- [12] "Problems That Fit", Numenta, Inc., cited Feb. 27, 2009  
<http://www.numenta.com/for-developers/education/ProblemsThatFitHTMs.pdf>.
- [13] "Hierarchical Temporal Memory – Comparison with Existing Models", Numenta, Inc., cited Feb. 27, 2009  
[http://www.numenta.com/for-developers/education/HTM\\_Comparison.pdf](http://www.numenta.com/for-developers/education/HTM_Comparison.pdf).
- [14] Pearl, Judea, (1985), "Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning". In *Proceedings of the 7th Conference of the Cognitive Science Society*, University of California, Irvine, CA, pp. 329-334, Aug. 15-17.
- [15] Ripley, Brian D., (1996), "Pattern Recognition and Neural Networks," Cambridge.