

# Identifying Style: Machine Recognition of Greek Temples and Gothic Cathedrals

Computational synthesis tools that automatically generate solutions to design problems are not widely used in architectural practice despite many years of research. This deficiency can be attributed, in part, to the difficulty of constructing robust building specific databases. New advances in artificial intelligence such as Hierarchical Temporal Memory (HTM) have the potential to make the construction of these databases more realistic in the near future. Based on an emerging theory of human neurological function, HTMs excel at ambiguous pattern recognition. This paper includes an experiment using HTMs for recognizing patterns in the form of visual style characteristics in Greek and Gothic architecture. Results from the experiment indicate that HTM is able to successfully categorize images of Greek and Gothic buildings. This is a promising development for future research in space and pattern recognition in architecture and the related design disciplines and suggests potential for this software to eventually be used as a generative tool within design automation efforts.

Architectural design computing has gone through a number of evolutions since its initial development in academia in the 1960's. The primary focus of many of the pioneers of architectural CAD computing was on the development of what is now known as knowledge-based computing software. The designers of early versions of CAD software saw the computer's potential as an intellectual companion to the architect, able to help with the more mundane tasks associated with the design process. These early CAD software prototypes included building specific databases that would aid in solving layout and code issues as the project was being developed. Despite early successes with knowledge-based software in academia these new forms of design computing never caught on in architectural practice, in part due to the difficulty of assembling building specific databases that were robust enough to truly be of service to the designer.<sup>1</sup> A second paradigm of CAD computing, evolving simultaneously, focused on the use of the computer as a digital replacement for hand drafting. The emergence of advanced graphic user interfaces, the evolution of graphic oriented input devices like the mouse, as well as advancements in printing capabilities, resulted in the second paradigm of CAD computing becoming the tool of choice in architectural

**JOSH HARTUNG**

University of Idaho

**FRANK R. JACOBUS**

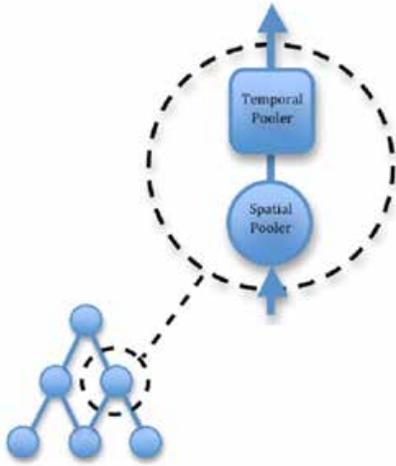
University of Arkansas

**GARRETT LUMENS**

University of Idaho

**JAY MCCORMACK**

University of Idaho



1

practice.<sup>1</sup> Evolution in computational speed and storage capacity over the past forty years, as well as current interest in newer forms of building information modeling (BIM), has caused many of the early knowledge-based CAD efforts to reenter the mainstream of architectural consciousness. A number of new knowledge-based computational tools are currently being explored and proposed for the architectural realm; the emergence of BIM tools is indicative of this next evolution. These tools promise a new paradigm in architectural practice in which our computer systems no longer simply act as unintelligent replacements for hand drafting, but instead become intelligent design companions that help guide architects through an increasingly complex design process.

The ability for these emergent design computation tools to perform effectively and therefore enter mainstream architectural practice relies on their continued evolution with respect to artificial intelligence and database management.<sup>2</sup> Within the Artificial Intelligence community 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,<sup>3</sup> case-based reasoning,<sup>4</sup> and rule-based systems such as shape grammars<sup>4</sup> use the capability to match a source example to target problems to reuse knowledge or determine the applicability of a rule. A new form of artificial intelligence that excels where some of the aforementioned systems have failed is Hierarchical Temporal Memory (HTM). HTM's are a form of artificial intelligence which are more proficient than previous systems at ambiguous pattern recognition. Based on an emerging theory of human neurological function, they can solve problems that traditional computing systems find difficult or impossible.<sup>5</sup> The potential for HTM's to recognize objects (specifically in the architectural realm the recognition of plan, diagram, form, wall, door, etc.) foretells the vast possibilities for how it may aid in the process of database building and object recognition within a design setting. This new intelligence could have a tremendous impact on architectural software and consequently the design process itself. This study includes a first experiment using chair back profiles to test HTMs for their ability to learn and recognize object shape and quality and compares the results to an identical test given to human subjects. Based on the results of the experiment an inquiry is made as to the potential impacts of HTMs on architectural practice and production.

### HTM SYSTEMS

Hierarchical Temporal Memory is a new type of artificial intelligence developed by Numenta, Inc. whose algorithmic structure emulates the structure of the human neocortex.<sup>6</sup> One of the primary advantages of HTMs is that they excel at ambiguous pattern recognition, which means they have the potential to solve problems which conventional computing systems cannot. HTMs have yielded positive results in traditional ambiguous problem areas within the computational realm such as semantic text analysis, fraud detection, and machine vision.<sup>7</sup>

Unlike traditional programming methodologies HTMs have both a training and an inference mode and are trained on input data. During the training mode the HTM analyzes an image to find specific patterns and shape information in space. These patterns and shapes are arranged in a hierarchy (Figure 1) based on their relationships in space and time. By identifying vector information and sending it through a spatial pooler the HTM can compare similar fragments of visual information in the form of vectors. Each unique vector is categorized by the HTM as a "coincidence;" a coincidence is a unique shape within the context of the overall image. If in this process the identified vector is similar to a vector that has

Figure 1: Temporal and Spatial Pooler.

already been through the spatial pooler then the HTM labels it as another coincidence of the exemplar vector. If the HTM does not recognize the vector then it is put into a separate category and considered a new coincidence. Once this process is complete for all vectors in a single image the unique categories are then sent to a temporal pooler (Figure 2).

The temporal pooler identifies patterns in time within the given image and receives category data from the spatial pooler in the sequence that the spatial data was identified. The sequence data is then evaluated by the HTM for coincidences, similar to the spatial pooler evaluation described above. Once the temporal coincidences are identified and categorized they are then disseminated up the nodal hierarchy. Each level of the hierarchy represents higher levels of complexity of object recognition. By the time the information reaches the top most nodes in the hierarchy the HTM has an invariant understanding of the problem data as a unified whole. This final categorization, which is comprised of all information from the lower and intermediate nodes in the hierarchy, is then used for Inference. The process of Inference is an attempt by the HTM to identify novel data from the same category as the training data. A remarkable attribute of the HTM which sets it apart from many previous pattern recognition systems is its ability to infer particular matches from highly ambiguous data.

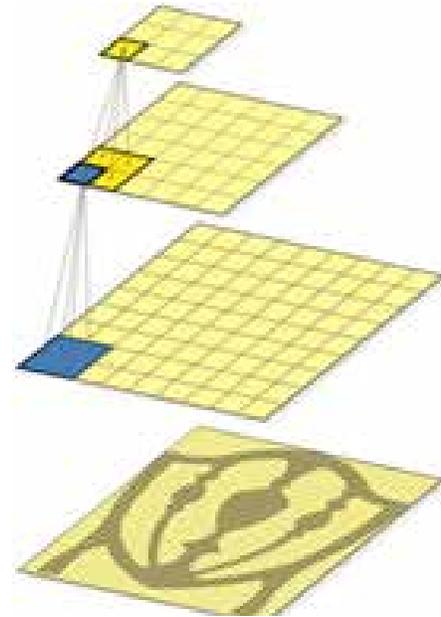
For example, if one were to train an HTM on images of a cow's entire body in a standing position and then run inference on a series of images that show the cow lying down or with its body partially obscured behind a barn, the HTM would still be able to recognize the figure in the image as "cow". Because the HTM still recognizes numerous spatio-temporal pattern matches between the training data and the novel image it is able to identify "cow" through probabilistic analysis. The system assumes, just as a human would, that if the patterns that make up the front of the cow are present in the image, then the patterns that make up the rear of the cow must be present but obscured by another object in the image.

During the process of inference the HTM draws from all levels of the hierarchy, not just the higher level nodes. This allows the HTM to store the underlying patterns that represent cows in all situations from a limited set of input data containing cows in limited situations. The categorized patterns form a model of quality that the HTM uses to define "cowness." Through the process described above the HTM stores underlying information about patterns that exist in space and time similar to the storage method of the human neocortex.

Though HTMs have many similarities to previously developed Neural and Bayesian networks, they differ in their ability to recognize and store spatially consistent data temporally within a hierarchy.<sup>8</sup> This distinction allows the HTM to build a robust and constant model of the input data and compute with high levels of ambiguity; a characteristic that most Neural and Bayesian networks do not share.<sup>9,10</sup>

## DEFINING STYLE

Before moving forward into the discussion of the experiment it is necessary to clarify what the word "style" means with respect to architecture. Movements in architecture occur over many decades, if not centuries, and it is often hindsight that allows us to look back at an epoch and categorize a group of buildings into a particular style. So then what mechanisms do we use to determine the properties of an era's style?



2

Figure 2: Nodal Hierarchy.

According to Witold Rybczynski, “architecture Style... describes the way that something is built,”<sup>14</sup> such as the specific use of form, techniques, materials, time period, and location. Moreover, “The theory of style... considers beauty as a unity, as a product or a result, not as a sum or a series. It seeks the constituent parts of form that are not form itself, but the idea, the force, the task, and the means, in other words, the basic preconditions of form.”<sup>15</sup> Buildings are classified into styles with few and broad principles, always with a handful of exceptions. Historians have defined these principles for the architectural styles. Every known factor that affected a building process or result is analyzed with comparison to these predetermined principles and the building is placed as best it can be into its particular category. Since “the history of architecture, including the present, is a continuity rather than a series of episodes”<sup>17</sup> this process can present challenges and disagreements. With this process being based on human opinions there will always be variability and discrepancy in how buildings are being classified. Technology has no personal opinion and thus would be able to make objective decisions fully supported by the rules and principles defined for each specific style based purely on the specific engineering training methodology of the identification program.

#### **THE CHAIR BACK EXPERIMENT**

In a previous experiment we explored HTM’s ability to recognize style and shape in existing chair back designs; we then compared the results to that of human beings tested on the same images. Three different chair designers were chosen for the experiment, Charles and Henry Greene, George Hepplewhite, and Gustave Stickley, and their chair backs were simplified to silhouettes. The HTM system was then trained and tested on a variety of chairs from each designer. A group of students were then trained and tested using identical images. We validated our study by structuring the human survey in a way that emulated the function of the computational model.

When trained and tested on identical chair back data the architecture students performed at the same level as the HTM. Both the students and the HTM tool placed the same Greene and Greene chair back into the Stickley category. Similarly, all but one student placed the same Stickley chair back into the Greene and Greene category, which is where the HTM placed it as well.

It became evident through analyzing the HTM and human inference decisions that the human and HTM models are very similar. These findings provide a strong basis for future experimentation with HTMs regarding design evaluation and synthesis. Relying on the chair back test as a foundation for the abilities of HTM to identify and classify similarly to human beings, we wanted to see if the system could recognize particular building styles. With this goal in mind we developed an HTM experiment in the recognition of Greek and Gothic architecture

#### **GREEK AND GOTHIC RECOGNITION EXPERIMENT**

The goal of the Greek and Gothic classification experiment was to see if HTM technology could identify Greek Temples as a distinct typology from Gothic Cathedrals by being trained on images from each respective type. Greek and Gothic were chosen as they have strong identifying, and in many respects contrasting, features which place them in their respective categories. For example, Gothic architecture often consists of pointed arches, rose windows, horizontal tripartite division, and has an overall vertical form. This vertical nature could either be on one side, both sides, or centered. Often times these heightened spaces had steeply pitched roofs, but this is not always the case. Notice, the test

was not trained to tell the difference between a lancet, equilateral, or flamboyant arch as this test is not yet developed to this level of detail. Greek is defined by a very consistent column structure, pediment base, shallow sloped roof, strong horizontal emphasis and a tripartite division along the y-axis. Greek architecture has a strong attention to exact symmetry while Gothic architecture often seeks for visual balance by the use of asymmetry. There are more distinct differences in the k architecture such as Greek architecture’s predominantly smooth marble or stone material usage contrasting with the fine detail and texture associated with Gothic architecture. Although not used as the main predominant identifiers, these features could help the HTM software in its identification process much like these features help the human eye differentiate.



3

Testing extreme cases of stylistic difference allowed us to train the HTM on fewer images due to the relative lack of nuance between the two styles. Success in this experiment would indicate likely future successes in experiments that attempt to recognize styles with minute variance in design types and categories. In establishing this test it was necessary to first identify what is constituted as “Greek” and “Gothic.” The test was not developed to recognize differences within a particular style that occurs as the style evolves or manifests itself in different geographic locations, nor to identify such nuances as the era the building was constructed. With this in mind, the general identifying features that a human would use to identify Gothic and Greek styled architecture were used and highlighted in this test. These features, as mentioned above, were created based on what most people would generalize as Gothic or Greek. These generalities are already established and can be seen by performing an internet search through any search engine for “Gothic Cathedral” or “Greek Temple.” The images that come up will be a variety of eras and locations (Figure 3) but will constitute a sense of what this test used as defining features of each style.

**HTM TRAINING PROCESS**

In order to effectively conduct the experiment it was necessary to be selective about the images used for training the system. By limiting the test and training images to front elevations, with only slight perspective, we were able to eliminate much of the potential confusion that vanishing lines might cause to the image as a whole. Eventually, with extensively populated image databases, through wiki-based collection methods, the recognition failures based on the perspectival problem likely will be eliminated. The training images were further refined to ensure that the HTM was focused on appropriate classification areas by removing everything in the background and limiting context to just the building in question. Likewise, the images were all cropped to the same size to ensure that the image edge had no effect on testing. Once these images were prepared

Figure 3: Internet Search of Gothic Cathedral.

they were uploaded into the Numenta Vision Toolkit. Approximately 100 cropped façade images were used for training for each building type (Figure 4). Once uploaded, the software automatically converted the images to gray scale. 30 original, façade images were then uploaded for testing.



4

Once the images were uploaded, the Numenta Vision Toolkit examined the training images for consistencies amongst the set. Following this, the software classified its belief distribution for how each image compared to the consistencies it found in the Greek and Gothic training images. Figure 5 is an example of the belief distribution provided by the HTM software after testing. Each category is tested independently, thus the total percentage need not equal zero. Each bar shown at the top right of Figure 5 indicates out of a possible 100 percent recognition of how closely the image relates to the training consistencies.

#### RESULTS OF THE EXPERIMENT

The software was trained on a total of 223 images, approximately 110 images for each category, and tested on 40 original images, 20 in each category. From this

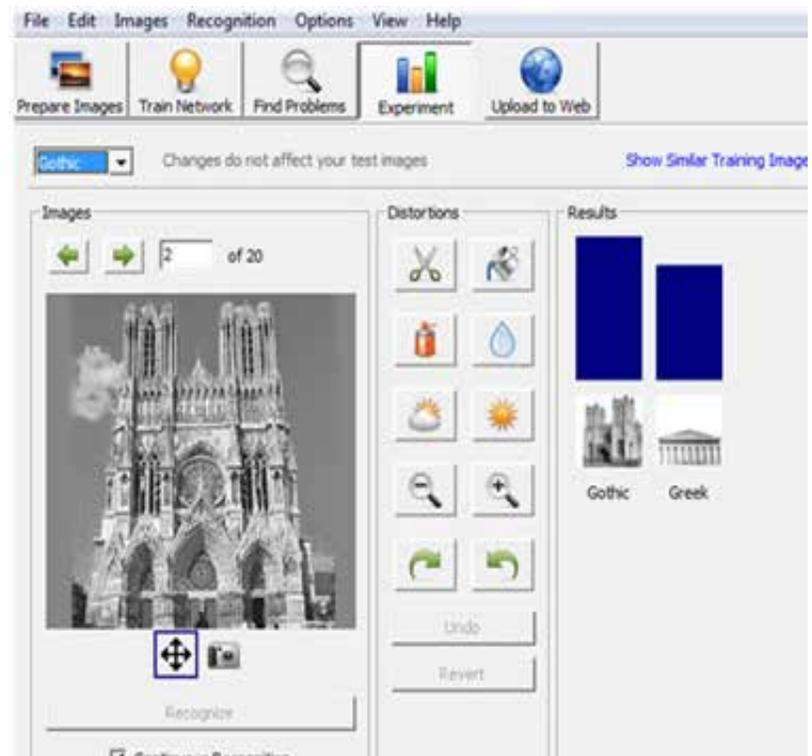


Figure 4: Selection of Training / Testing Images.

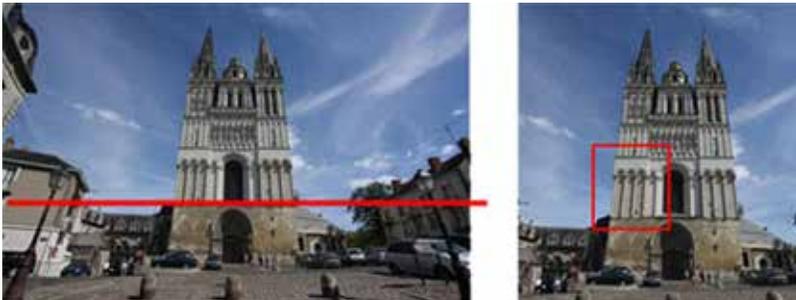
Figure 5: Inference on Rheims Gothic Cathedral.

5

testing the HTM tool was able to place all 40 of the images back into the appropriate category which constitutes a 100% success rate. A test with all of the same training data and testing images, where the testing images weren't cropped to the same size showed a 90% success rate. This is believed to have occurred because the edge of the image would create a false identification line. Once the images were cropped to the same dimensions, the 100% accuracy was attained. The failure of the other method helps us to understand the limitations of the HTM software. Unlike a human that is able to account for the pictures being different sizes and shapes, the HTM software needs a definitive set of variables which it is analyzing to be able to accurately categorize the testing images.

### BACKGROUND CONFUSION

Unlike the human eye, the HTM software does not know how to separate the background from the foreground or focus of the image. This ability is what allows humans to identify a specific building without distraction by its surroundings or context. HTM's inability to separate the background means that the background will affect and play a part into the image identification. As can be seen in Figure 6, the background creates a strong horizontal line, consistent with the Greek Temple rather than the Gothic Cathedral. The presence of other physical features in the image causes the HTM tool to identify shapes that should not be associated with the Gothic Cathedral itself, thus resulting in an inaccurate identification of the image in question. When this image was cropped to the standardized square dimension, the majority of the background was lost, allowing the software to accurately identify the building typology.



6

### LACKING AN ATTENTION MECHANISM

Despite HTMs abilities to accurately identify the two architectural styles, it typically has a low level of confidence in its selection. Conversely, the human brain likely would have a strong confidence level in identifying two styles that are so radically different. For the HTM software this could be just a matter of needing more training data to increase the confidence level as a human has a lifetime of training images and data to use for identification. This is because the Numenta software lacks an attention mechanism, and so is not a fully realized model of human neurological function. Since both buildings are comprised of vertical lines, horizontal lines, and diagonal lines, there are natural formal similarities between them. HTM is identifying both stylistic types as a series of shapes which form what we humans know to be a building. A portion of the confidence that HTM produces in comparing the test images to the training images is from parts of the image that the human will naturally disregard as unimportant. The small variability in results that shows the building is Gothic over Greek, or vice-versa, is all the human brain is focused on analyzing.

Figure 6: Image Identification Issues.

When a human compares the two objects, they are able to ignore the similarities that they are both buildings and identify the building independent of its background and surrounding context and look specifically for the stylistic differences that they're trying to identify. The differences are weighted heavily and thus are more influential to the human brain than they are to the HTM software.

The inverse problem exists as well. HTM recognizes patterns and systems at all scales. As highlighted in Figure 7, Angers Gothic Cathedral has what visually appears to be a series of columns. The human eye would be able to see that these features are not structural but instead only decorative features. Thus a human identifier would weight this feature lower in classifying the building. HTM Software cannot detect what is decorative instead of structural. Moreover, humans recognize that a Greek Colonnade continues for the full length of width of the façade. It is clear in this image that the column pattern is broken. HTM would not recognize this break as it lacks an attention mechanism.

### **QUALITY EVALUATION**

The success of this test shows potential for this system to be further developed and potentially even start to classify era, geographic, or other such style nuance variances. was based upon the aforementioned limitations, including standardized image sizing

Despite the success with these two contrasting styles it is unclear if this method could be further developed and potentially start to classify era, geographic, or other such style nuance variances with further training. It must be mentioned that the success of this test was based upon the aforementioned limitations, including standardized image sizing and a façade only test. But as a starting point with time limitations, the success this test achieved shows future development possibilities. It's possible that if the HTM was trained with perspective images and a wider berth of reference images that the software could identify more challenging images and angles of the building. This test with limited training data was able to accurately identify two separate architectural styles and acts as a great starting point showing that HTM and similar software have a potential future in architecture and it's probably safe to say that with a more inclusive and extensive training set HTM would be able to recognize more nuances.

### **FUTURE USES**

This first step in computers categorizing architectural styles lays the foundation for further development and investigation for building identification databases to help with research, building spatial and patterning characteristics as well as design automation efforts. This test led to findings in the current level of Hierarchical Temporal Memory technology:

- HTM can affectively recognize trained stylistic patterns in standardized testing images.
- HTM cannot identify an image background as distinct from the image focus and thus identifies based on the background context as well.
- An attention mechanism is vital for the further development of HTM software.
- The more training data developed and used the better and more accurately an image is identified.
- Color images (converted to gray scale) can be used successfully for testing.

Concurrent with this experiment, other tests related to the architectural realm were exploring HTM's ability to recognize individual designers and its capacity for plan and diagram recognition related to building type, all with a relatively high degree of success. These tests indicate that with further refinement and development of HTM and other artificial intelligences they can be directly applied to the realm of architecture, much like they already have been to speech and character identification.

This test's training was limited to the front elevation of the building. The training data depth could be further developed and expanded to recognize particular parts of the building by being trained specifically on pointed arches, column capitals, etc. Training to this level of detail would start to divide the classification systems into their own categories such as Doric, Ionic, and Corinthian or Lancet, Equilateral, or Flamboyant. With a further development of the parts of each building which classifies it as that building type, it would be possible for the system to be further developed to include more style types and expand the classification system a great deal. However, before we spend all our efforts trying to reach too far forward it must be recognized that this software can be used for other tasks with little further development.

#### **FUTURE RESEARCH AND EDUCATIONAL USES**

This sort of identification software could help establish a standard classification system for past and present architectural projects. With the high amount of research that is now done on the internet, the information found is only as accurate as the person uploading it was knowledgeable. Since the internet is an open source it is inconsistent and has no self-checking system to ensure reliability or accuracy. If an image is uploaded with an incorrect title or classification, the unaware researcher could easily be misled. If HTM technology was included in standard practice of internet uploading, a wider array of images could be accurately classified and research could thus be more easily completed. Before this sort of identification system would be successful it would be necessary for knowledgeable individuals to train the systems during its initial development stages. The time and effort for this training would initially need to be exerted, but once established, the system would be able to continue to expand its own training data by successfully identified test images. With this method of self expanding, once established even if the system was to inaccurately identify a few images and add them to the training data, their features would be overwhelmed and simply by nature of quantity of examples the inaccurate image would become an outlier.

Studies have shown that most people learn more by doing than by any other means of study. If an identification system was included in educational uses, it would be possible for assignments to be given to students to design a building that met a certain stylistic classification. Students would then be directly interacting with previous design styles while still having the ability to expand beyond preexisting styles. Even if for sake of curiosity it would be possible for students to use an identification system to categorize their buildings after they have finished their designs. This future of design classification could be spread to all other design disciplines including, but not limited to, interior, graphic, product, and furniture design.

#### **FUTURE DESIGN USES**

The use of HTM identification systems would not be limited to the educational world. In the professional world clients often come to a firm or an architect with a general notion of a design in mind. With a robust enough identification library it

would be possible to identify the overall style they were aiming to achieve. More specifically related to space design, it would be possible to take existing spaces with certain spatial qualities and relationships and help to generate an environment based on one which was previously categorized, rather than simply recreate a specific style classification.

As the digital age evolves and computers become a stronger design ally it is possible that rather than HTM software being merely an identification tool, it could be a creation mechanism as well. As systems are further developed that devise and create an overall building parti or organization, HTM could be coupled into the system to generate an overall style or classification that they want the building to invoke. Once features, rhythms, patterns, and visual languages have been established for a classification style, the computer would be able to create new building appearances that fit the 'rules' established for that style. The idea doesn't have to be limited to Gothic Cathedral vs. Greek Temple, but could instead start to include residential style categories such as prairie style, post modernism, etc. A rough project outline and visual appearance could be generated almost instantaneously for the client to see and either reject or like as a starting point for further development.

The use of HTM technology is one possible future of architecture that would allow computers to influence the design, research, and classification of all buildings. The limitations of how HTM could be used are only limited based on what level of development has been input into the level of sophistication of the HTM software itself, the creation of the training data, and the mind that is looking for ways to use it. Development and use of systems such as HTM are necessary as the world of architecture continues to change and connect to the coming paradigm that will be further indebted to computational tools.

## ENDNOTES

1. Kalay, Y. E. *Architecture's New Media*. The MIT Press, Cambridge, Massachusetts, 2004.
2. Novák, V., Perfilieva, I., and Močkoř, J., *Mathematical Principles of Fuzzy Logic*, Kluwer Academic, Dodrecht, 1999.
3. Leake, D., *CBR in Context: The Present and Future, In Case-Based Reasoning: Experiences, Lessons, and Future Directions*. AAAI Press/MIT Press, pp. 1-30, 1996.
4. Stiny, G., Introduction to *Shape and Shape Grammars, Environment and Planning B: Planning and Design*, Vol. 7, pp. 343-351, 1980.
5. Hawkins, J., *On Intelligence*, Times Books, 2004.
6. 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)>
7. Hitchcock, Henry-Russell and Johnson, Philip, *The International Style*, W. W. Norton & Company, Inc. New York, NY, 1966.
8. Problems That Fit, Numenta, Inc., cited Feb. 27, 2009 <http://www.numenta.com/for-developers/education/ProblemsThatFitHTMs.pdf>.
9. 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).
10. Pearl, J., "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, 1985.
11. Ripley, B.D., *Pattern Recognition and Neural Networks*, Cambridge, 1996.
12. McCormack, J., Dr., Hartung J., Jacobus, F., "Support for the Use of Hierarchical Temporal Memory Systems in Automated Design Evaluation: A First Experiment." Proceedings of the ASME. Proceedings of the American Society of Mechanical Engineers IDECT/CIE. San Diego, CA: American Society of Mechanical Engineers, 2009.
13. Muthesius, Hermann, *Style-Architecture and Building-Art: Transformations of Architecture in the Nineteenth Century and its Present Condition*, The Getty Center for the History of art and Humanities, Santa Monica, CA 1994.
14. Rybczynski, Witold, *The Look of Architecture*, Oxford University Press, New York, NY, 2001.
15. Semper, Gottfried, *The Four Elements of Architecture*, Cambridge University Press, New York, NY, 1989.