18 Connectionist CAAD for Restructuring Shapes in Terms of Emergent Subshapes

Yu-Tung Liu

Architecture Group, Graduate School of Applied Arts, National Chiao Tung University, Hsinchu, Taiwan

Designers naturally restructure shapes in terms of emergent subshapes in the process of design. According to the result of a psychological experiment about how experienced and non-experienced designers see shapes, only experienced designers can encode implicit subshapes emerged from the primary shapes. Many symbolic approaches have been considered in addressing this focused problem. On the other hand, the issue is also encountered by connectionist networks, also called parallel distributed models or neural networks. Recognizing both explicit and implicit emergent subshapes has been explored using connectionist networks associated with appropriate mechanisms of visual attention, namely recurrent attention and searchlight attention in combination. The distinction between symbolic and connectionist computations of shapes is discussed.

INTRODUCTION

Both design knowledge and design concepts draw extensively upon the representations of shapes. Designers naturally restructure shapes in terms of emergent subshapes to aid their further deliberations in design. It has been generally accepted that the recognition of emergent subshapes plays an important role in design derivation [xxiv, xxv, xxvi, xxxiii, xxxiv, xxxv]. The ability to vividly encode subshapes can thus be viewed as a touchstone for the development of computer-aided architectural design (CAAD) systems. Many symbolic approaches have been considered in addressing this focused problem [xxvii, xxxvi, xxxv, viii]. Symbolic processing, without doubt, provides some advantages such as computational efficiency, systematicity, and productivity [xxi,vii]; however, it is still limited to neatly reducing the solutions to a manageable, significant portion and to dealing with noisy and unexpected input information [xxi, vi], as an experienced designer does. In addition, the issue of shape recognition can be explored by connectionist networks which essentially attempt to model human brain functions by connecting vast artificial neurons in different ways. It is commonly known that visual recognition is an activity of the brain, specifically the activity in the visual area of the cerebral cortex [ix]. The study of restructuring shapes, which is still limited from a symbolic point of view, may go deeper from a connectionist point of view.

In design search, design evolves from one state to another by exhaustively or heuristically applying proper rules. Each rule application involves, first, pattern-matching the antecedent of a rule to the current state and, second, transforming the matched portion of that state into the consequence of the rule. However pattern-matching techniques of current CAAD systems are still limited. For instance, consider the shape in Figure 1a which is drawn by two squares. In current CAAD systems, only those two squares can be dealt with by pattern-matching for further development as shown in Figure 1b. However, even a non-experienced designer can effortlessly recognize not only those two but other explicit emergent subshapes, for example a smaller square in the middle where the two squares overlap and two L-shapes in the corners as shown in Figure 1c. In addition to those explicit subshapes, an experienced designer is able to encode some implicit subshapes, such as an implicitly emerging square in the lower-left corner as shown in Figure 1d [xviii]. Therefore a human designer can thoroughly deliberate all these alternatives before making a decision. In other words, a human designer can restructure shapes in terms of emergent subshapes at any step in the design process.

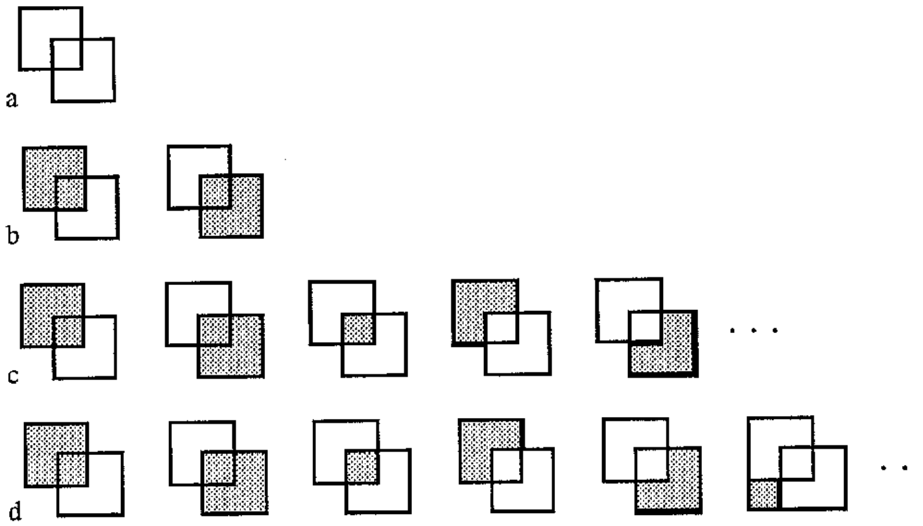


Figure 1. Recognizing two overlapping squares by (b) current CAAD systems, (c) non-experienced designers, and (d) experienced designers

The objective of this paper is to encode both explicit and implicit emergent subshapes using connectionist networks associated with some mechanisms of visual attention. The distinction between symbolic and connectionist processing to be used in CAAD will be discussed in due course.

1.0 BACKGROUND ON CONNECTIONIST CAAD

In the area of CAAD, two-layered connectionist models were previously adopted to solve building attribute/form mapping problems [iii, iv]. Nevertheless, because two-layered models lack of the hidden layer's internal representation, Minsky and Papert [xxiii] have declared that two-layered models are deficient and thus inappropriate to be used for problem-solving. Consequently, a multi-layered model is constructed that provides capable problem-solving for building attribute/form mapping problems [xv]. In human design processes, many drawings of shapes remain incomplete or are executed inaccurately, namely irregular and "noisy" shapes. The recognition and transformation of those ill-processed shapes have been adequately explored by multi-layered BackProp networks [xv]. To simulate emergent subshape recognition [xvi], a training set for input and expected output patterns, consisting of 25 shapes in different scales, locations, and rotation angles, are used as prototypes of human visual experience about shapes (Figure 2). The bitmap representation of the 25 training shapes and their corresponding object-oriented representations were utilized when various networks were constructed for different simulations. After training the network to recognize the 25 shapes, a resulting set of weights in a well-learned network was referred to as a *solution set of weights (SSW)*.

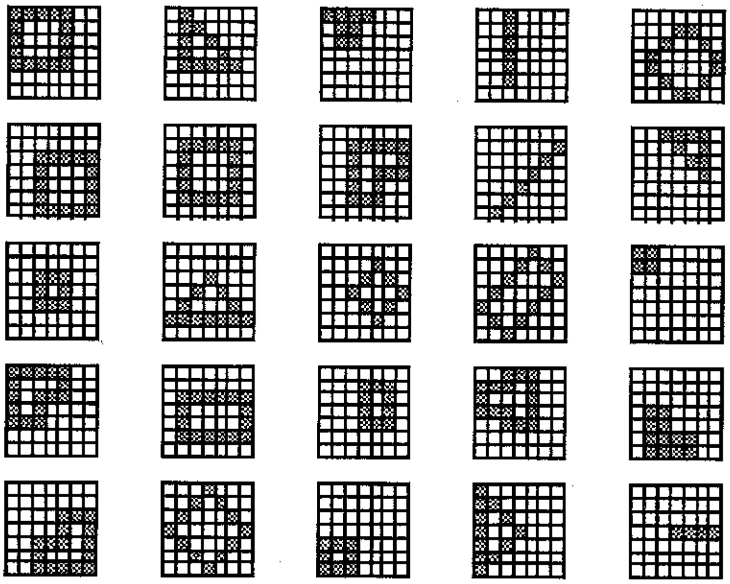


Figure 2. Training patterns for the BackProp network's learning procedures

By virtue of the multi layer connectionist model's great capability for input-output mapping, when a well-learned network "sees" any accurate shape (Figure 3a) or even an ill-processed one (Figure 3b) as one of the testing input patterns, its exact, complete corresponding shape can be acquired respectively. The results are coincident with the simulations on ill-processed shape recognition [xv]. Nevertheless, the behavior of the BackProp model is much more sophisticated when the testing input pattern contains multiple shapes in one "retinal" grid. As it presents two separate shapes to the well-trained network as shown in Figure 3c, the resulting output is a distracted image where the two testing shapes are both weakly activated but neither dominates. When the two overlapping squares indicated in Figure 1a are presented, it can be seen that the output pattern is also distracted: many subshapes are only partially or weakly activated, as shown in Figure 3d; again, none dominates. To manipulate multiple shapes at the same time, a regular multi-layer network cannot differentiate them as expected without an attentional technique .

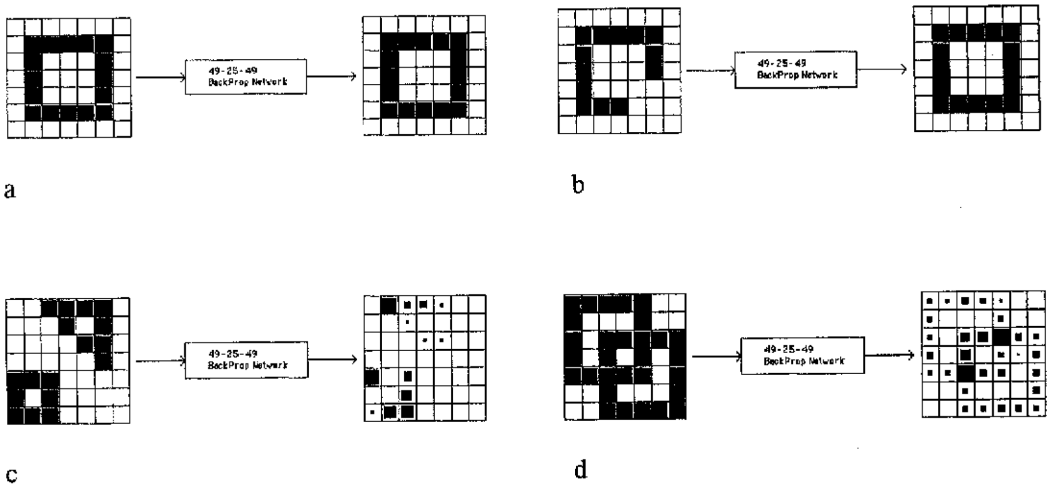


Figure 3. Testing inputs containing (a) an accurate square; (b) an ill-processed square; (c) a triangle and a s quare; and (d) two overlapping squares and their corresponding bitmap outputs

To address this problem as the information limit on pattern recognition for both human beings and computers [xx(422-423)], recurrent attention has been proposed; this means, technically, putting the resulting distracted output from a well-learned network back into the input layer of the same network (Figure 4). The underlying logic is that although many shapes are partially activated together, one of them is either slightly or distinctly stronger than the others. By using the resulting output recurrently as the next input, the stronger shape gets stronger and the weaker ones get weaker. This process of attention is repeated until the stronger shape becomes overwhelmingly active and all the others are eliminated. Using recurrent attention, each SSW is guaranteed to find a primary shape or an emergent subshape. Because each learning procedure starts with a random initial set of weights, each solution set of weights should be distinct and thus potentiate recognizing different subshapes. By acquiring ten SSWs in the training processes and filtering out repetitive subshapes, the system can obtain the five explicit, closed emergent subshapes [xvi].

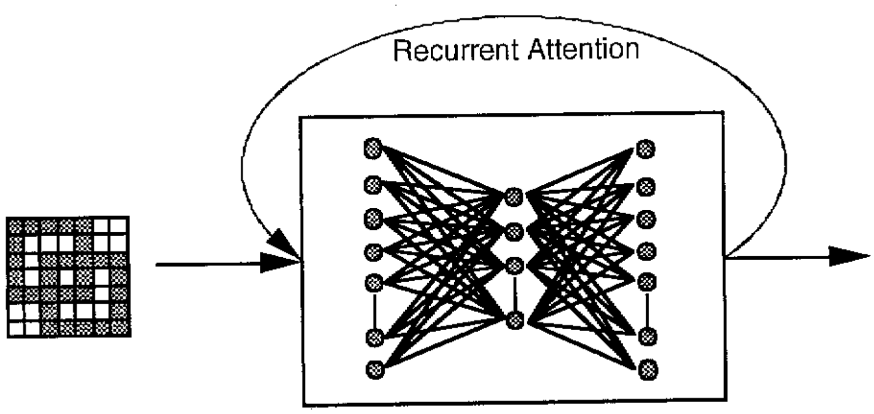


Figure 4. The multi-layered network with recurrent attention

Knowledge of shapes represented in a bitmap fashion can be applied immediately to bitmapped, pixel-based systems. But the attended subshapes in the bitmap representation are still inappropriate to integrate with many other drafting and CAAD systems because these systems draw on symbolic representations of knowledge and manipulate design knowledge in a symbolic way [xiv]. By employing another connectionist network encoding bitmapped input and objected-oriented output, an attended subshape acquired in the bitmap fashion can be converted into an object-oriented form and further into its symbolic representation as *(Square 3 3 3 3 0)*, as illustrated in Figure 5.

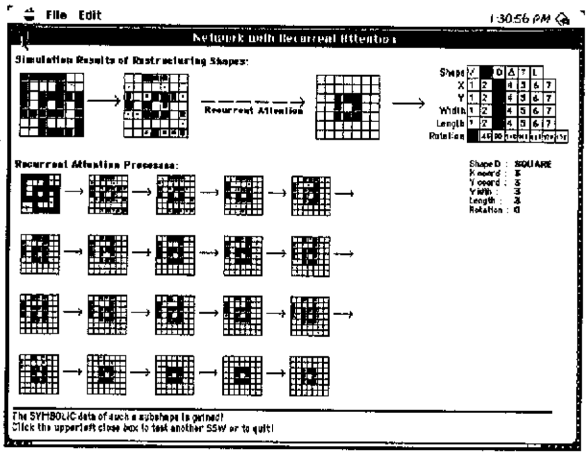


Figure 5. A demonstration of the entire procedure for restructuring shapes based on one SSW

Although the procedure for restructuring shapes using connectionist networks and recurrent attention can solve the problem appropriately, it still has some limitations. First of all, the logic of recurrent attention is that when multiple shapes are presented concurrently, some subshapes are insubstantially activated, and one of them is stronger than the rest. After paying attention for a longer time, the stronger shape will overwhelmingly dominate. However, when the scales of primary shapes are extremely different, such a method no longer works; the recurrent attention always ends up on the largest shape. Second, to implement the mechanism of recurrent attention requires a number of training processes to get numerous SSWs and each SSW guarantees finding a subshape. Therefore, a different mechanism of attention which is able to encode multiple emergent subshapes in only *one* training process is an important direction for future work. Third, as demonstrated in an empirical study of seeing shapes [xviii], although implicit subshapes are rarely found by non-experienced designers, they are common for experienced designers. Another procedure for restructuring shapes is needed in the search for implicit shape recognition.

2.0 SEARCHLIGHT ATTENTION PROCESSING

The concept of recurrent attention [xvi] is that within a specific field of vision which includes multiple objects, people naturally and gradually focus on an object by paying more attention recurrently to the stronger stimulus. That concept addresses the emergent subshapes problem to some degree, but encounters some limitations, as outlined above. Drawing upon the extensive findings on the role of searchlight attention in visual performance, this study proposes a searchlight attention mechanism to be embedded in the connectionist networks, bearing on the same problem again in order to overcome some of the limitations of the recurrent attention concept.

Based upon Neisser's [xxviii] notion of the distinction between pre-attentive and attentive processing, the distracted images gained from the output layer of a connectionist network without any attentional mechanism (Figure 3c, d) can be seen as pre-attentive processing. The typical characteristic of such a period is that it is a holistic operation of objects, of which none can be identified. Paying attention to the object one sees requires the viewer to select a portion of the current visual field. This visual search task has been described as the "searchlight" of attention: A searchlight moving in the scene of a visual field by adjusting the locations of the searchlight as well as the positions of the eyes, head, and body as needed [v, ix, xiii]. Now, in order to get beyond this information limit, the searchlight attention process strives to encode emergent subshapes in the visual field one at a time. To mimic the phenomenon of searchlight attention, some critical variables and parameters must be defined first, drawing from the psychological findings on searchlight attention. Obviously, the first variable is the searchlight *locations* which shift around the entire receptive field; the 7-by-7 "retinal" grid used in this study is one example. The second variable to be implemented is the searchlight *scopes* which vary from 4-by-4 to 6-by-6 squares in this study. Note that there is no evidence that the aperture of the searchlight is necessarily square or circular; the 4-by-4, 5-by-5, and 6-by-6 square apertures are arbitrarily chosen for the sake of simplicity. In sum, not only does a searchlight move throughout the presenting primary shape in a right-to-left and down-to-up fashion, but also its scope is adjusted systematically, as shown in Figure 6.

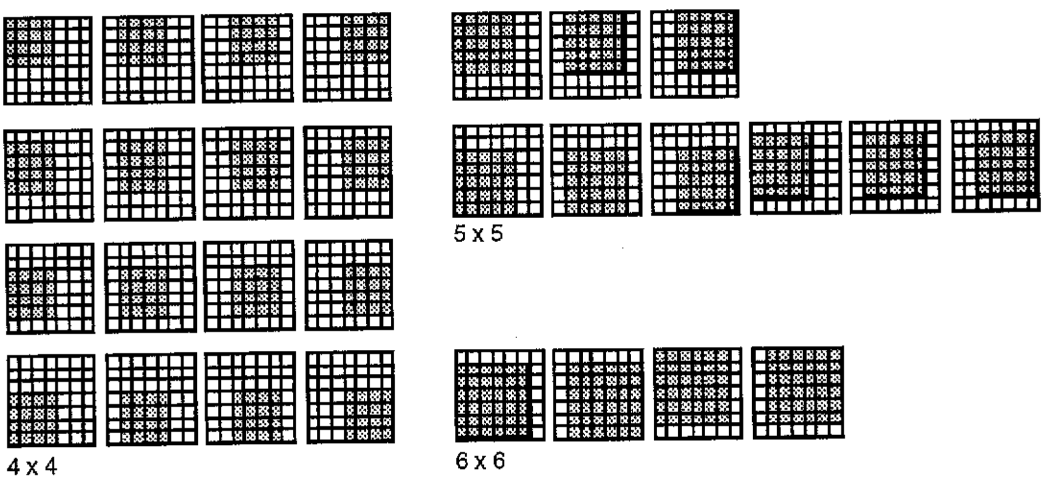


Figure 6 The locations and scopes of a searchlight of attention on a 7-by-7 "retinal" grid as a receptive field (The scopes, 4-by-4, 5-by-5, and 6-by-6, are arbitrarily chosen)

In a sense, the searchlight attention process that can see shapes is a peculiar *problem decomposition*: Since the problem of encoding multiple subshapes simultaneously is too difficult for people to solve, in our visual system the problem is naturally decomposed, or articulated, into parts, each of which includes only a portion of the information in the entire visual field [xx]. That is, by means of the searchlight attention mechanism, an insoluble problem is decomposed into solvable subproblems, in this case the 29 attending processes shown in Figure 6. Additionally, decomposing a problem into subproblems is often regarded as a significant strategy for bearing on ill-defined problems I, [xxxii]. Here it implies that, in essence, problem decomposition is not only a strategic tools for solving hard problems but also a necessary means for overcoming the information limit, for both humans and computers.

Depending on the threshold recognition paradigm proposed by Poirson and Wandell [xxxi], an attaching threshold is critical to distinguish between the target and the distracted image, not only for people but for computers. A threshold idea of this kind has been previously implemented in constructing the extracting algorithm in the recurrent attention mechanism for processing both bitmapped and object-oriented outputs of shapes. The threshold adopted in the recurrent attention mechanism, named the threshold of recognizing activation (TRA), is 0.9 accordingly; it is adjustable here in an attempt to mimic the suggestion that ordinary people attach high thresholds, while experienced designers have lowered their thresholds in order to see shapes more flexibly [xviii]. In addition, when people shift attention and discern a specific scope in a visual field, the image outside the scope can also be seen weakly. From our visual experience, different people should see an image outside the scope of attention at different strength, under different situations. Therefore, it is plausible to add another variable, called activation strength for the outside searchlight attention.

3.0 COMBINED PROCESSING: SEARCHLIGHT AND RECURRENT ATTENTION

Two connectionist networks and two mechanisms of selective attention are embedded in pursuit of better attentional performance for recognizing emergent subshapes. As illustrated in Figure 7, a primary input shape is first processed by the searchlight attention mechanism; in this way, the whole task of recognition is then decomposed into attending portions. Each searchlight-attended input is further dealt with by a bitmap-in-bitmap-out network in association with the recurrent attention that presumably focuses on only one shape within the current searchlight window. Furthermore, the final attending image can be processed by a bitmap-in-object-oriented-out network in order to get the object-oriented data about that shape. Finally, by applying an extracting mechanism to the final output of shape, the object-oriented data is thus converted into a symbolic form, for the systems integration reason as mentioned earlier.

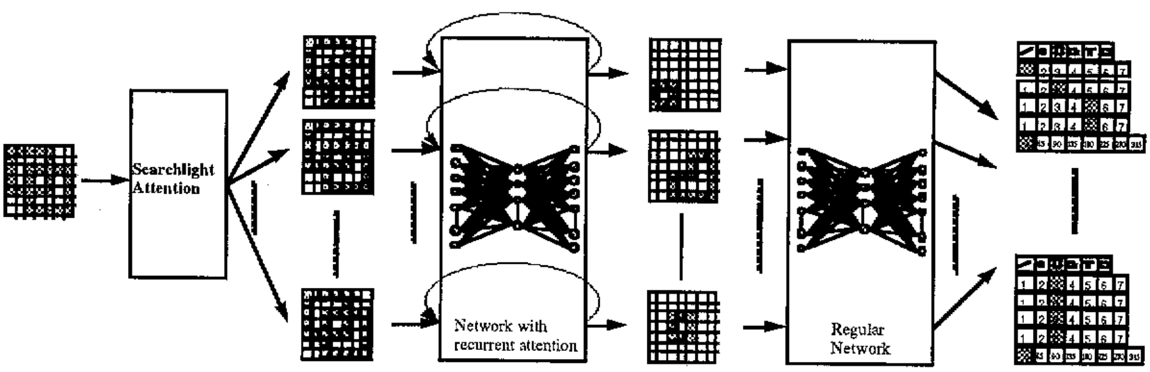


Figure 7. The procedure for restructuring shapes using the searchlight and recurrent attention mechanisms in combination, before the processes of the extracting mechanism and STM storing

The above procedure relies effectively on three variables—TRA, the activation strength for the outside searchlight, and the number of cycles for the recurrent attention. The focus here is the overall network behavior. The following simulation is based on a primary situation: TRA is set at 0.75, the outside activation strength at 0.13, and the number of recurrent cycles at 1. Note that in the previous simulations based on recurrent attention alone, the number of recurrent cycles ranged from 10 to 20, adapting to different complexities in the presenting shapes. The many recurrent cycles were needed because an initially stronger subshape had to compete with a number of weak subshapes within the entire visual field, a 7-by-7 "retinal" grid. Here, prior to the start of the recurrent attention processing, the searchlight attention mechanism has already excluded many of the weak subshapes around the stronger one. As a result, the number of recurrent cycles here is efficiently reduced to 1 which is enough to handle the complexity of the primary shape, the two overlapping squares.

The simulation results are shown in Figure 8 in which the five explicit, closed subshapes are successfully encoded when the searchlight moves to adequate locations in the visual field. Note that only one training process is required for each network to achieve the above visual performance, compared to the many training processes needed to execute the process using individual recurrent attention.

As discussed above, the third and the most challenging limitation to be overcome lies in the network's ability to recognize the implicit subshape. Based on the current setting of variables for the simulation (Figure 8), even using two attentional mechanisms in combination it is still unable to find the implicitly emergent subshape, namely the lower-left smaller square. The phenomenon of experienced and non-experienced subjects seeing shapes [xviii], suggests that experienced designers have lowered their naturally attached TRA values, which enables them to see some implicit subshapes emerging from the primary shapes. This concept provides a clue to recognizing the implicit subshape throughout the procedure proposed here.

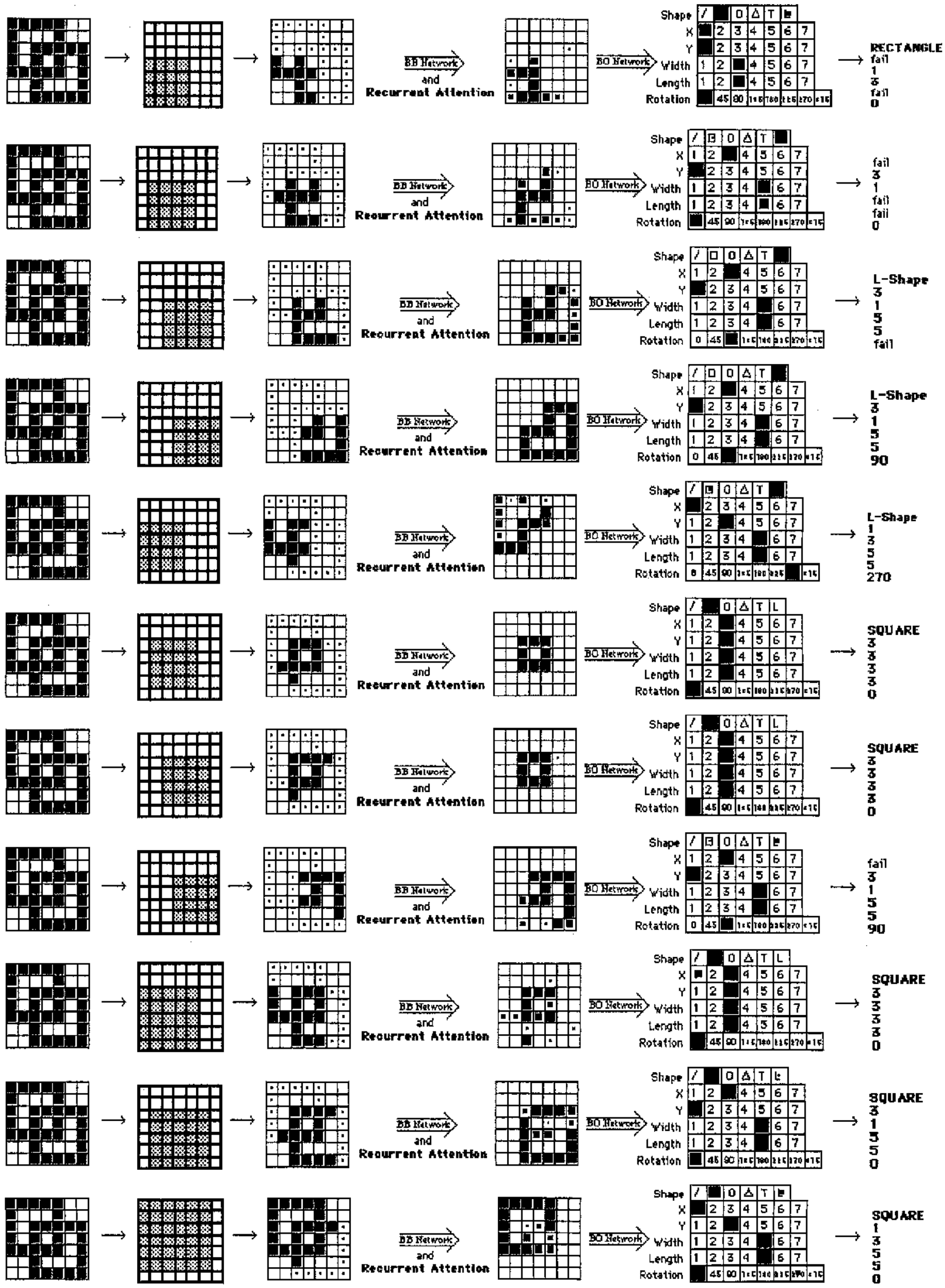


Figure 8. The simulation results of encoding two overlaid squares using the network with both the searchlight and recurrent attention mechanisms. (Only some locations and scopes of the searchlight are shown; other unsuccessful attending processes are not included)

As shown in the first example in Figure 8, when the two attentional mechanisms are applied to the primary shape, the final attending image in a bitmap fashion is close to a square; therefore the corresponding object-oriented data can be correctly extracted, but the activations are not strong enough to exceed the threshold of 0.75 (also shown in Figure 9a). This is why it fails in that simulation. After lowering the TRA variable from 0.75 to 0.65, the same procedure successfully encodes and extracts the implicit subshape in a symbolic form, as illustrated in Figure 9b. Decreasing the TRA, however, is not the only method that can reveal the implicit subshapes. As in the first procedure in Figure 9, the final attending image is still too distracted to be a square after the two attentional techniques are used. This is because, within the specific "lit-up" portion of the field, the close-to-square image is still not overwhelmingly dominant when only one cycle of recurrent attention is applied to it, as pre-defined. It is plausible to suspect that the close-to-square shape would become more dominant after more than one recurrent attention cycle and that it would successfully exceed the original TRA of 0.75. Such an idea is proved and illustrated in Figure 9c.

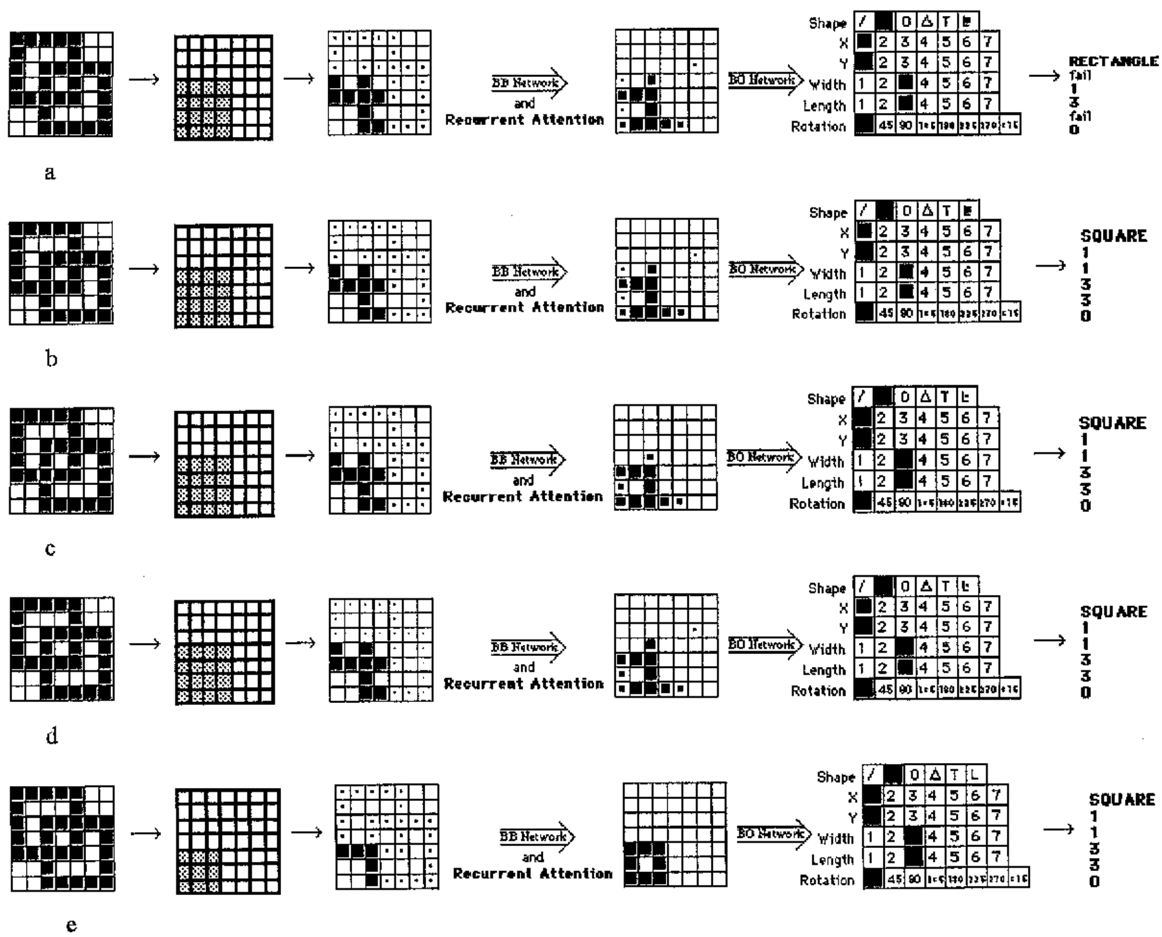


Figure 9. Simulation result of encoding the implicit subshape by (a) using the original variable setting; (b) decreasing the TRA from 0.75 to 0.65; (c) slightly increasing the number of recurrent cycles to 2; (d) decreasing the activation strength for the outside searchlight from 0.13 to 0; and (e) shifting a 3-by-3 aperture of the searchlight attention to an appropriate location

Another alternative is to decrease the activation strength for the outside searchlight. As discussed, when focusing on a specific portion of the receptive field, we can still feebly see images outside that portion. Although there is no psychological data about this, it is still plausible that when looking at a specific portion of the visual field, a person can see images outside the attending portion at different strengths, adapted to different viewing situations. Thus, decreasing the outside strength, that is, eliminating more of the interference from the outside objects, enables the networks to recognize some implicit subshapes. This hypothesis is validated as shown in Figure 9d.

Moreover, as explained above, the simulation is based on the arbitrarily-chosen 4-by-4, 5-by-5, and 6-by-6 apertures for searchlight attention. If the searchlight includes a 3-by-3 aperture moving around the visual field, the 3-by-3 implicit square could be easier to find by the procedure with the same TRA of 0.75. In other words, there is a most appropriate size for the searchlight for each specific shape in each specific location. The result of this concept is positive and shown in Figure 9e, when the searchlight moves to the critical position, adjusting the aperture to the critical size.

In summary, the limitation on recognizing implicit subshapes can be overcome by individually varying some critical variables, namely TRA, the number of recurrent cycles, and the size of the searchlight aperture. Slight adjustments in some variables allow the network to see an implicit subshape. This reveals not only how complex the visual mechanisms in our mind are, but also how sophisticated a computer program has to be in order to perform the same task as people do.

4.0 SYMBOLIC VERSUS CONNECTIONIST COMPUTATIONS

Throughout this paper on restructuring shapes in design, connectionist processing has been used to address the issue of emergent subshapes which traditionally are addressed by symbolic processing [xxxiii, xxxiv, xxxv, xxxvi, xxvii,, viii]. For these two opposing approaches, an advantage in one is clearly a shortcoming in the other, and vice versa. In the areas of AI and cognitive psychology, the two approaches are generally recognized as forms of antagonism rather than cooperation embedded in the human mind. For example, as mentioned earlier, Allen Newell [xxix(129)] claimed that neural processing is very limited in its ability to produce fully cognitive behavior which normally takes place in the order of seconds, because of the extremely quick operation time of neural networking which happens in tens of milliseconds (Table 1). On the other hand, connectionist advocates treat the so-called behaviors cognitive as simply the side effects of low-level neural networking [xix].

To the ongoing arguments between advocates of symbolism and connectionism, Marvin Minsky [xxi(35)] provides another point of view: "much has been said, in the popular press, as though these were conflicting activities. This seems exceedingly strange to me because both are parts of the same enterprise." Both approaches have different strengths and weaknesses because they underlie totally different philosophies: Symbolism believes that the human mind manipulates symbols and structures of symbols; thus it calls the mind's organization a symbol system or information processing system [xxxii]. On the other hand, connectionism relies on the assumption that the basic units posited in cognition are processed in a simple, uniform way—connecting neurons in specific ways.

Table 1: Time scale of human action, after Newell [xxix(122)]

Scale (sec)	Time Units	System	World (theory)

10^{7}			months
10^{6}	weeks		**Social Band**
10^{5}	days		
10^{4}	hours	Task	
10^{3}	10 min	Task	**Rational Band**
10^{2}	minutes	Task	
10^{1}	10 sec	Unit task	
10^{0}	1 sec	Operations	**Cognitive Band**
10^{-1}	100 ms	Deliberate act	
10^{-2}	10 ms	Neural circuit	
10^{-3}	1 ms	Neuron	**Biological Band**
10^{-4}	100 μs	Organelle	

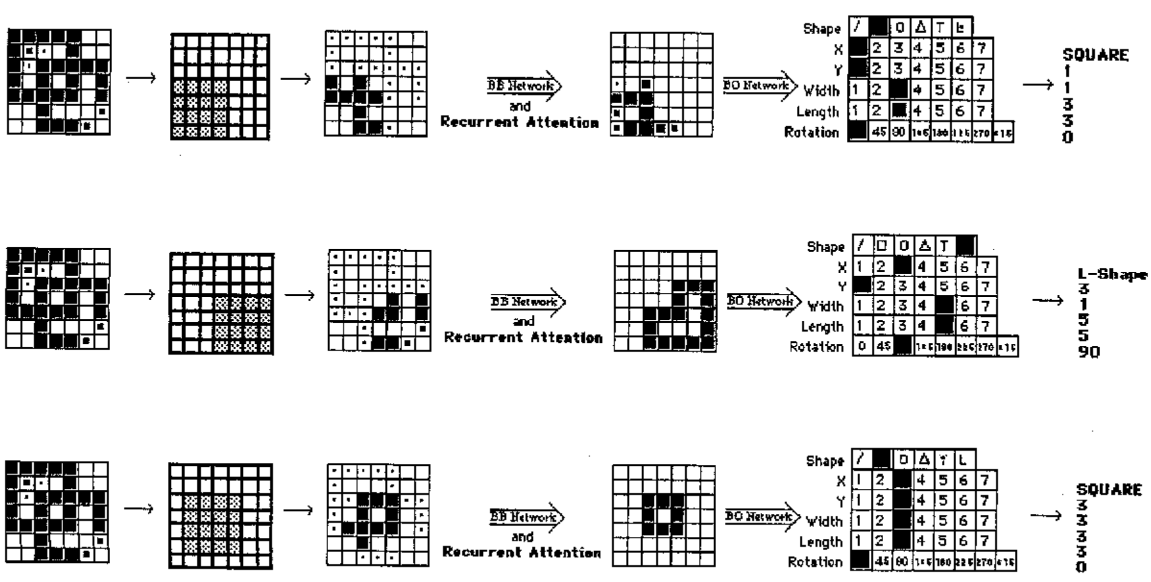
Consider the characteristics of the symbolic approach first. In addition to being capable of managing high-level cognitive behavior such as reasoning and search [xxi, vi], other general advantages of symbolic computation are its *efficiency*, *systematicity*, and *productivity* [vii] which are completely shown in symbolic solutions of the emergent subshape problem [xxxiii, xxxiv, xxxv, xxxvi, xxvii, viii]. Practically speaking, efficient computing is always a worthy direction to pursue. However, productivity may not always be viewed as a positive, depending on how many alternatives the system produces. In terms of the classification of emergent subshapes including the explicit and implicit, and closed and unclosed shapes defined in the previous section, the number of emergent subshapes could be infinite. Of course, the symbolic, algorithmic approach can efficiently produce a great number of solutions. Simon [xxxii(66)] sees problem-solving as a search: "successful problem solving involves searching the maze selectively and reducing it to manageable proportions." By analogy, a successful attempt at emergent subshape recognition should produce a manageable proportion out of all possible emergent subshapes, those subshapes significantly important to most people and to designers.

As generally pointed out, the common disadvantages of the symbolic approach belong to its inflexibility in handling different problem domains and its inability to deal with noisy and unexpected input information [xxi, vi]. Any symbolic system, once built, is only good in the domain as pre-defined; it cannot adapt the domain-specific symbolic module to another problem domain. Moreover, symbolic systems draw on well-structured representations of symbols and semantic attributes associated with them to generate results. But, the symbols that human beings commonly deal with are not always well-structured; they are normally in incomplete, irregular, and noisy forms. The symbolic computation of shapes is inherently limited in abilities to handle these two issues well.

In contrast, the connectionist model seems to provide natural solutions to many of the serious problems that arise from the symbolic model. First of all, although connectionist networks are not highly productive, the number of results a network can generate is directly based upon the number of associative patterns it has learned. Practically speaking, by carefully controlling the kind and number of training patterns, the results of a connectionist network can be restricted to a manageable portion to be further processed. Second, in numerous connectionist simulations in AI and cognitive psychology, the uniform structure of the multilayer connectionist network has been suitable for every low-level cognitive task, such as memory and recognition. In other words, the connectionist approach is highly adaptable to various problem domains. Furthermore, a previous connectionist study [xv] illustrated the network's robust ability to handle ill-structured representations of shapes, and to restructure ill-processed shapes into regular forms. That was simply the case of isolated shape simulations. How well can the network handle multiple, ill-processed shapes?

In a new simulation, the test scheme includes two overlapping squares of which one is incomplete and the other is noisy due to some unexpected dots as shown in Figure 10. This ill-structured "symbol," which is extremely difficult for the symbolic model to handle, here is manipulated by the procedure including two networks associated with the searchlight and recurrent attention processes (Figure 7). The simulation result indicates that the networks are also able to successfully restructure the ill-processed image by generating both the explicitly emerging subshapes and the implicitly emerging one, as shown in Figure 10.

So far, all discussions in regard to connectionism are positive. This does not imply that parallel distributed processing is excellent for simulating every human cognitive behavior. Actually, the connectionist model has two serious limitations, namely systematicity and processing efficiency, which are mentioned to as major strengths of symbolic processing. The lack of systematicity [7] here means that connectionism does not allow for the *decomposition* and *recomposition* of representations of knowledge, which are ubiquitous and important in human cognition and can be manipulated well by symbolic computation. In addition, the lack of efficiency in connectionist processing is obvious: in this study for example, not only is the training process extremely time-consuming, but also the shape restructuring procedure takes a serious amount of time, if we compare the computation needed to process numerous connections in the two networks and that to serially process knowledge in symbolic systems.



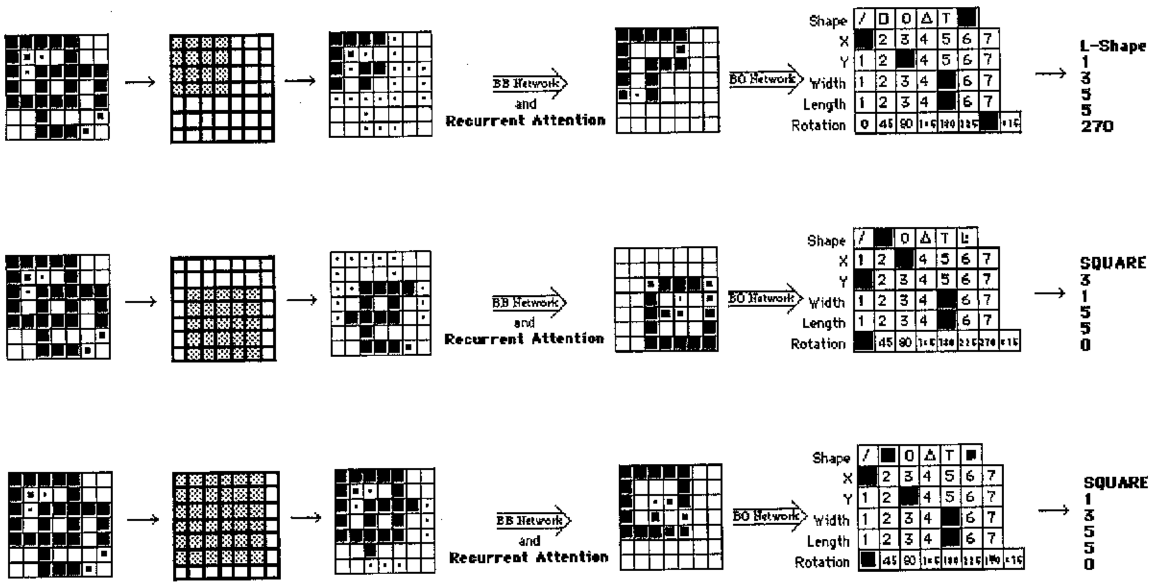


Figure 10. The simulation results of encoding two overlaid, ill-processed squares using the network with both the searchlight and recurrent attentions (See Figure 8 for note)

Given the above comparisons, it would be unreasonable to claim that one approach is better than the other. In actuality, both encode different levels of knowledge and process that knowledge in different ways to pursue different, necessary kinds of performance. Based on the investigations of AI, cognitive psychology, and the study of design shape in this paper, it seems to be true that the mostly serial, symbolic computation is good at high-level cognition such as reasoning and search; the parallel distributed connectionist computation, in contrast, is good at low-level cognition such as pattern recognition and memory retrieval. A sound CAD system needs both. This is why the shape restructuring procedure in this study includes a symbolic extracting technique rather than a connectionist network. For the same reason, emphasis is placed on systems integration, through conversion from the bitmapped into the symbolic representation of shapes. The most challenging question, in AI, cognitive psychology and design, is how to use the two apparently conflicting approaches in significant ways, as a whole, to get better performances and a deeper understanding of human cognition (Figure 11).

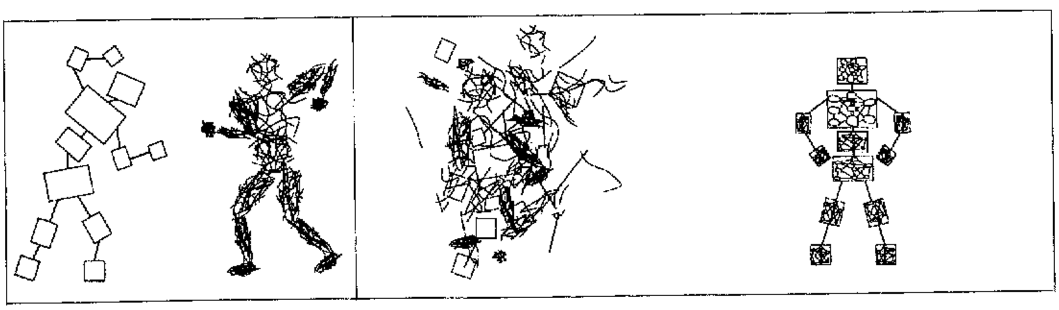


Figure 11. Conflict between theoretical extremes, after Minsky [21]

To integrate connectionist models into symbolic CAAD is by no means easy. The major obstacle is that, at present, connectionist and symbolic approaches belong to two opposite *logics*. First, the primitives in the connectionist models, neurons, are very small in scale and numerous in quantity; whereas those in the symbolic systems, symbols, are comparatively very large but their numbers are much smaller (Figure 12). Second, symbolic systems are processed serially; while connectionist models are processed in parallel. These two distinctions make it difficult to integrate these two approaches and even to communicate between them. Thus, exploring the intermediate area between connectionist and symbolic processing is the inevitable, primary mission for the next decade in AI and cognitive psychology [xxii] as well as in CAAD.

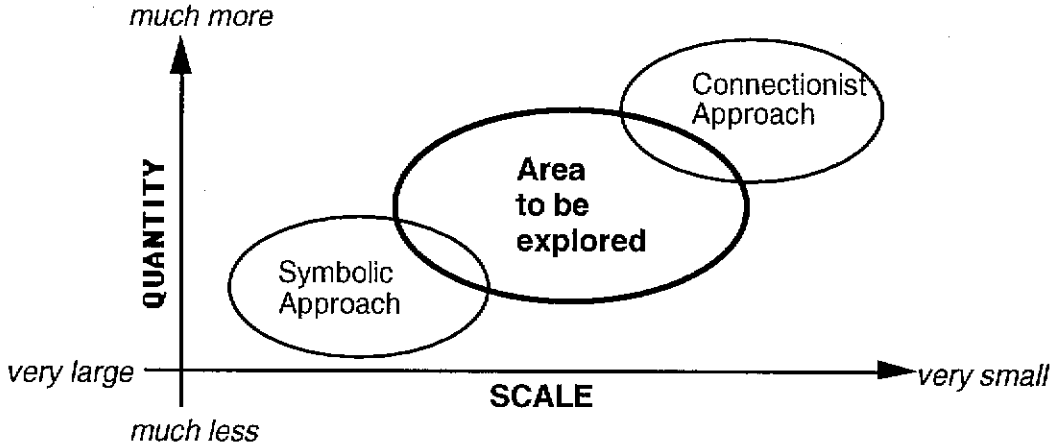


Figure 12. *Symbolic approach versus connectionist approach in scale and quantity, after Minsky [22]*

CONCLUSION: INTERACTION BETWEEN COMPUTATION AND COGNITION

In summary, this paper started by building a model of seeing emergent subshapes based on some characterization and phenomena of human visual behavior. After a structure was proposed for the model, a computer system using connectionist networks associated with attentional mechanisms was then constructed in order to encode both the explicit and the implicit subshapes that emerge from the primary shapes. Then, the computer outcome was used to validate the proposed model by comparing the computer data and the empirical data under some specific situations. As the result of this process, I found two important points about this CAD application. First, the understanding of a designer's cognition is extremely helpful in constructing a CAD system and in overcoming some of the limitations it might face. Second, the behavior of the CAD system capable of encoding emergent subshapes provides a straightforward, solid way to validate the proposed model of restructuring shapes. In other words, the interaction between design computation and design cognition is important not only for CAD studies but also for cognitive studies of design (Figure 13).

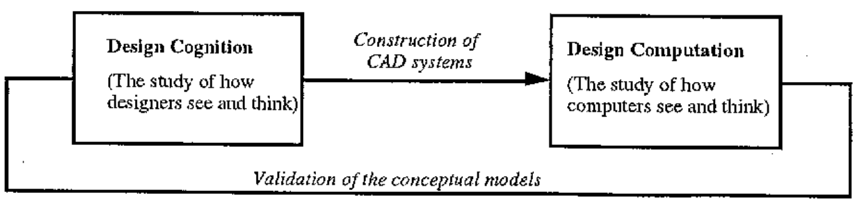
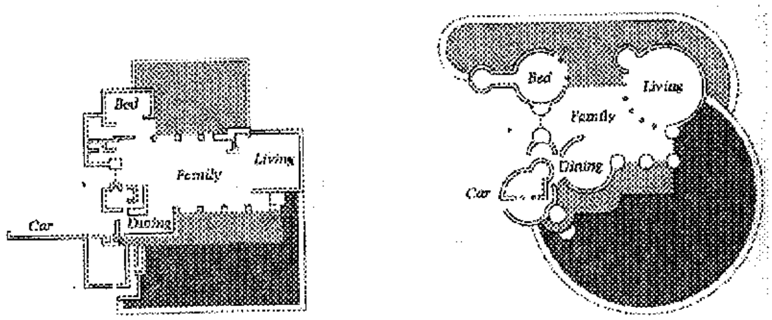


Figure 13. The interaction of design cognition and design computation

Technically, the connectionist network used in this study is still in its infancy; therefore, although it shows some advantages, there are a number of limitations. Some of these are global problems to be encountered in psychology, neuroscience, and AI, while others are truly local problems of design. The following domain-specific limitations to be solved in design form the future directions for this research:

Emergent subshape preference. One future study lies in determining how we select "types of essential subshapes" out of the great number of subshapes that emerge from the primary shapes. For example, if the square is the essential element in a house plan design, the square emergent subshapes should be more important than all the others (Figure 14a), while circular emergent subshapes should be first encoded in a design whose fundamental primitives are circles (Figure 14b).



a (Life House, 1938) b (Jester House, 1938)

Figure 14. Some Frank Lloyd Wright house plans

Non-dimensional scheme for the receptive field. In this study, a 7-by-7 receptive field is used as the simulation basis to represent any shape in a bitmapped fashion. Future studies should include a zoom-in/zoom-out system for networks so that drawings in different scales can be processed by fitting them into the predefined receptive field. In other words, the networks should be able to deal with non-dimensional drawings for practical purposes.

REFERENCES:

[i] Ö. Akin, *Psychology of Architectural Design* (Pion, London, 1986).

[ii] D.E. Broadbent, Stimulus set and response set: Two kinds of selective attention, in: D.I. Mostofsky, ed., *Attention: Contemporary theory and analysis* (Meredith, New York, 1970) 51-60.

[iii] R.D. Coyne, Modelling the emergence of design descriptions across schemata, *Environment and Planning B: Planning and Design* 18 (1991) 427-458.

[iv] R.D. Coyne & S. Newton, Design reasoning by association, *Environment and Planning B: Planning and Design* 17 (1990) 39-56.

[v] F.H.C. Crick, The function of the thalamic reticular complex: The searchlight hypothesis, in: proc. of *The National Academy of Science*, USA, 81, (1984) 4586-4590.

[vi] J. Dinsmore, Thunder in the gap, in: J. Dinsmore, ed., *Symbolic and connectionist paradigms: Closing the gap* (Erlbaum, Hillsdale, NJ, 1992) 1-23.

[vii] J. Fodor & Z.W. Pylyshyn, Connectionism and cognitive architecture: A critical analysis, in: S. Pinker and J. Mehler, eds., *Connections and symbols* (MIT Press, Cambridge, MA, 1988) 3-72.

[viii] J.S. Gero & M. Yan, Discovering emergent shapes using a data-driven symbolic model, in: U. Flemming & S. Van Wyk, eds. *CAAD Futures '93* (Elsevier, Amsterdam, 1993) 3-17.

[ix] R.L. Gregory, ed., *The Oxford Companion to the Mind* (Oxford University Press, Oxford, 1987).

[x] L. Henderson, *Orthography and Word Recognition in Reading* (Academic Press, New York, 1982).

[xi] W. James, *The Principles of Psychology* (Dover, New York, [1890] 1946).

[xii] K. Koffka, *Principles of Gestalt Psychology* (Harcourt, Brace and Company, New York, 1935).

[xiii] S.M. Kosslyn & O. Koeing, *Wet Mind* (Free Press, New York, 1992).

[xiv] Y.T. Liu, Schematic-Designer: A knowledge-based CAD system for schematic design in architecture, *Design Studies* 12(3) (1991) 151-167.

[xv] Y.T. Liu, Restructuring shapes in terms of emergent subshapes: A computational and cognitive model. Accepted to appear in *Environment and Planning B: Planning and Design*.

[xvi] Y.T. Liu, Symbolic CAAD versus Connectionist CAAD. *Languages of design* 2(4) (1994) 325-346.

[xvii] Y.T. Liu, Encoding explicit and implicit emergent subshapes based on empirical findings about human vision, in: J.S. Gero & F. Sudweeks, eds., *Artificial Intelligence in Design '94* (Kluwer, Dordrecht, Netherlands, 1994) 401-418.

[xviii] Y.T. Liu, Some phenomena of seeing shapes in design. *Design Studies* 16(3) (1995) 367-385.

[xix] J.L. McClelland, .D.E. Rumelhart & G.E. Hinton, The appeal of parallel distributed processing, in: D.E. Rumelhart, J.L. McClelland & the PDP Research Group, eds., *Parallel distributed processing* (MIT Press, Cambridge, MA, 1986) 3-44.

[xx] M. Minsky, Steps toward artificial intelligence, in: E.A. Feigenbaum & J. Feldman, eds., *Computers and Thought* (McGraw-Hill, New York, 1963) 406-450.

[xxi] M. Minsky, Logical versus analogical or symbolic versus connectionist or neat versus scruffy, *AI Magazine* Summer (1991) 34-51.

[xxii] M. Minsky, Lecture on *The Society of Mind*, Media laboratory, MIT (March 4, 1992).

[xxiii] M. Minsky & S. Papert, *Perceptrons* (MIT Press, Cambridge, MA, 1969).

[xxiv] W.J. Mitchell, *The Logic of Architecture* (MIT Press, Cambridge, MA, 1990).

[xxv] W.J. Mitchell, The uses of inconsistency in design, in: Y.E. Kalay, ed., *Principles of Computer-aided Design: Evaluating and Predicting Design Performance* (John Wiley, New York, 1992) 1-13.

[xxvi] W.J. Mitchell, A computational view of design creativity, in: J.S. Gero & M.L. Maher eds., *Modelling Creativity and Knowledge-based Creative Design* (Erlbaum, Hillsdale, NJ, 1992).

[xxvii] T. Nagakura, Shape recognition and transformation: A script-based approach, in: M. McCullough, W.J. Mitchell, and P. Purcell, eds., *The Eectronic Design Studio* (MIT Press, Cambridge, MA, 1990) 149-171.

[xxviii] U. Neisser, *Cognitive Psychology* (Meredith, New York, 1967).

[xxix] A. Newell, *Unified Theories of Cognition* (Harvard University Press, Cambridge, MA, 1990).

[xxx] J. Palmer, C.T. Ames & D.T. Lindsey, Measuring the effect of attention on simple visual search, *Journal of Experimental Psychology: Human Perception and Performance* 19(1) (1993) 108-130.

[xxxi] A.B. Poirson, & B.A. Wandell, Task-dependent color discrimination, *Journal of the Optical Society* 7 (1990) 776-782.

[xxxii] H.A. Simon, *The Sciences of the Artificial,* 2d ed. (MIT Press, Cambridge, MA, 1981).

[xxxiii] G. Stiny, What designer do that computers should, in: M. McCullough, W.J. Mitchell, and P. Purcell, eds., *The Eectronic Design Studio* (MIT Press, Cambridge, MA, 1990) 17-30.

[xxxiv] G. Stiny, Weights, *Environment and planning B: Planning and Design* 19 (1992) 413-430.

[xxxv] G. Stiny, Emergence and Continuity in shape grammars, in U. Flemming & S. Van Wyk, eds. *CAAD Futures '93* (Elsevier, Amsterdam, 1993) 37-54.

[xxxvi] M. Tan, Closing in on an open problem-reasons and a strategy to encode emergent subshapes, in proc. of *ACADIA '90 Proceedings: From research to practice*, Big Sky, Montana, (1990) 5-19.