Action Selection Under Constraints: Dynamic Optimization of Behavior in Machines and Humans

Robert Kozma, Derek Harter and Srinivas Achunala
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152, U.S.A.
email: {rkozma, dharter, sachunal}@memphis.edu

Abstract - Biological brains are capable of adaptive behavior to sustain performance in tasks in the face of increasingly difficult constraints. We present a task with varying conditions of resource and time constraints. We compare our heuristic and neural network models to human data and speculate about dynamic mechanisms of action selection.

I. INTRODUCTION

The fundamental question for all biological organisms comes down to what should be done next [1]. In the study of autonomous agents this has come to be known as the **action selection** problem. Biological brains not only solve this problem well, but show amazing abilities to adapt to changing constraints in the task and environment. In other words, biological organisms are able to dynamically adjust to changing constraints and maintain good performance on tasks in the face of increasing difficulties.

Biological brains are fundamentally pattern-forming, self-organizing systems governed by nonlinear dynamical laws [2, pg. 26]. It has been shown that nonlinear, chaotic dynamics are used in the formation of perceptual categories in biological brains [3], [4]. We believe that such dynamics are not only essential in the formation of perceptual meaning, but also in the formation of a shifting hierarchy of intentional goal states, that we observe as the action selection behavior of biological organisms [5], [2], [6], [7].

II. PACKING TASK TESTBED

To study the performance of action selection under constraints we have developed a packing task as shown in figure 1, which is a variant of the popular tetris computer game [8], [7], [9]. In our packing task, the subject is presented with a series of 10 blocks, that appear at the top of the playing field. There are 3 basic block shapes. In a sequence of 10 blocks which constitutes a single trial, the subject will receive different block types chosen at random. Blocks can be positioned by moving them left or right, or by rotating them clockwise or counter-clockwise. Once positioned by the subject they are dropped onto the playing field. When a block drops onto the playing field, it descends until it reaches the bottom or is obstructed in its downward fall by another block.

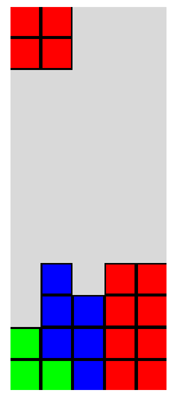

Fig. 1. The packing task. Blocks appear from the top and the subject rotates and moves the block before dropping it onto the playing field. The goal is to obtain as dense of a packing as possible.

The goal of the task for the subject is to pack the shapes into the bottom of the playing field as tightly as possible. The density of their packing, which is a measure of the subjects success on the task, can be calculated simply by dividing the area filled in with blocks with the total area. For example, in figure 1 the playing field currently has 5 columns with 4 rows in height for a total area of 20. Out of that area of 20, 17 cells are filled with blocks. Therefore in the figure, the current density of the packing is 17 / 20 or 0.85.

The task, simple as it might seem, is still too difficult to perform optimally for a human (3 different blocks can be placed in 28 orientations with 10 blocks per trial gives a search space of 28^{10} or $3\text{x}10^{14}$ possible sequences). Further the task is made more difficult by the introduction of constraints that must also be considered when choosing behavior. In this paper we discuss simulations of the packing task under two different types of constraints: resource and time.

Resources are constrained in the packing task by giving the subject only a certain number of translation/rotation resources at the beginning of a trial. For example, if the subject is given 15 resources, they will only be able to make a combination of 15 moves and rotations over the whole 10 block trial. Each move (left or right) and each rotation (clockwise or counter-clockwise) expends one of the subjects resources. When the subject runs out of resources before the end of a trial, any remaining shapes simply fall at random on the playing field.

Time constraints are the second type of constraint modeled in these simulations. When a subject is playing the packing task under a time constraint, they will be presented with a block at the top of the playing field and given only a certain amount of time to position the block, for example 1 second. When time runs out, the block falls whether the subject has finished placing it in their intended position and orientation or not.

Time and resource constraints further complicate the packing task and make optimal play impossible. Under such constraints, systems are forced to produce behavior in noisy conditions and under uncertain information. But even under such unfavorable conditions, biological systems are capable of maintaining performance levels in the face of increasingly difficult constraints.

In the next section we present a computer simulation of a behavior producing system performing under time and resource constraints in the packing task. In section IV we present some data collected on human subjects performing the packing task and compare it with our computer simulation. Finally we discuss the implications of our research for dynamical models of action selection in autonomous agents.

III. COMPUTER SIMULATION

A. Neural Network and Heuristic Algorithms

The action selection mechanism presented here is meant to model some aspects of biological organisms in producing behavior on the packing task under constraints. In particular, we model the selection of an intended goal position for a block using a neural network or a set of heuristics.

We have developed several algorithms and heuristics to perform the packing task based on various principles. Among these are neural networks based on backpropogation learning and a heuristic algorithm using heuristics derived from studying the behavior of people on the task. The neural network based model is illustrated in details in [8], [7]. It involves a multilayer perceptron trained on examples created by human experts. The neural network algorithm shows reasonable performance with top scores close to the those achieved by human experts. The results of the neural network and human experts are given in table I.

TABLE I
NEURAL NETWORK PERFORMANCE

	Density
Human	0.8748
Neural Network	0.8261

Our computer simulations using heuristics evaluate the resulting situation of dropping a given block in a particular orientation and position onto the current playing field. The heuristic evaluation takes into account factors such as the resulting contour shape of the playing field, and the creation of unfillable holes, as well as other features. These factors can be combined to evaluate the desirability of placing a block at a particular location and orientation given the current situation.

The heuristic evaluation of intended goal positions is used as a starting point in the decision making process. When performing the packing task under constraints, behavior may be modified when constraints are factored into the decision making process. For example, which move is considered *best* may be very different if there are plenty of resources left as opposed to when there are only very few resources left.

B. Expectancy and Resource Constraints

Resource constraints can influence the behavior producing mechanisms in biological organisms. People seem to be able to intuitively adjust their behavior on the packing task to improve performance and minimize problems from running out of resources. For example when people have plenty of resources they freely expend resources on good moves that may cost a lot of resources. However, people seem to switch strategies and will select less desirable moves that help conserve resources when they perceive they are running out of them.

We have modeled this intuitive conservation of resources under conditions of constraint using a factor we call *Expectancy*. Expectancy is a measure of the expected number of resources needed on average for each block in a trial. In this case it is the number of resources that are expected to be needed for each block in order to obtain a reasonably good packing performance. For example, suppose that you intuitively feel that you need 20 resources in order to pack 10 blocks reasonably well. Another way of looking at this would be that you expect to expend, on average, 2 resources for each block in order to obtain a good packing. Given this intuitive expectancy of 2 resources per block, you can dynamically alter your behavior during a trial in order to expend your resources wisely. If you have 5 blocks left to pack you would expect to need about 10 resources to obtain a good packing performance. If you actually have 15 resources left you would feel fairly safe in choosing the move you think best, even if you have to expend 5 resources in order to execute it. However, if you only had 8 resources instead, you might think twice about expending 5 resources on a move and instead pick a slightly worse move that helps to conserve resources.

Of course we don't believe that people consciously make such calculations while performing the packing task. However, they do intuitively develop something like an expectancy parameter through experience in performing the packing task. This intuitive feel of expected resource usage guides the subject in modifying behavior appropriately under various conditions of resource constraints that they encounter. Expectancy is simply an intuitive strategy that people develop with experience on the packing task. Experience on the packing task helps people to develop a natural expectation of the number of resources they usually need per block (or over the whole trial) in order to obtain good performance.

Our algorithm uses a comparison of expected resource usage to actual resources remaining as described above to help choose moves that balance between conservation of resources and optimal moves.

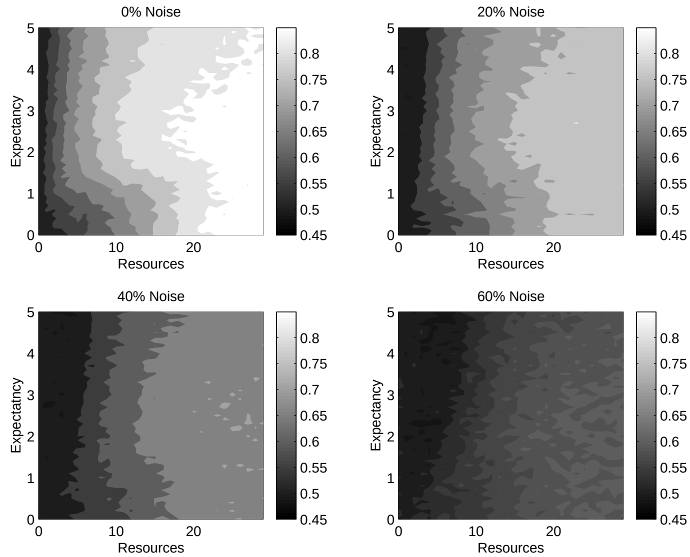

Fig. 2. The computer simulation of the packing task using heuristics. This figure shows the average density achieved by the algorithm at 0,20,40 and 60% noise levels.

C. Noise and Time Constraints

The second type of constraint modeled in these simulations is constraints on time. Time pressure can be added to the packing task by limiting the amount of time given the subject, from presentation of the block to when the block falls, for the subject to perform moves and rotations in order to place the block in their intended goal position. Time constraints manifest themselves as pressures to act. As time pressures are increased, behaviors may change that favor easy moves that are less prone to error and confusion and that can be accomplished more quickly. For example, rotations are a much more difficult manipulation to perform compared to translations, and much more prone to errors. As time pressures are increased, human players rely less and less on rotations and favor translation manipulations.

Time constraints manifest themselves in human performance in various ways, but the ultimate effect is to induce an error. By an error, we mean that the subject fails to move and position the block to their intended goal location. This may happen because they run out of time before they complete their sequence to the intended location, or time pressures may increase the likelihood of producing an unintended behavior.

We model time constraints in our simulations by introducing noise, or random errors, into the simulations. One example model of error production is to say that some percentage of the time the block does not end up in its intended goal location, but instead ends up in some other location at random. A more realistic model is to simulate the sequence of moves needed to transition from the initial location to the intended goal location. In the more realistic model, each move

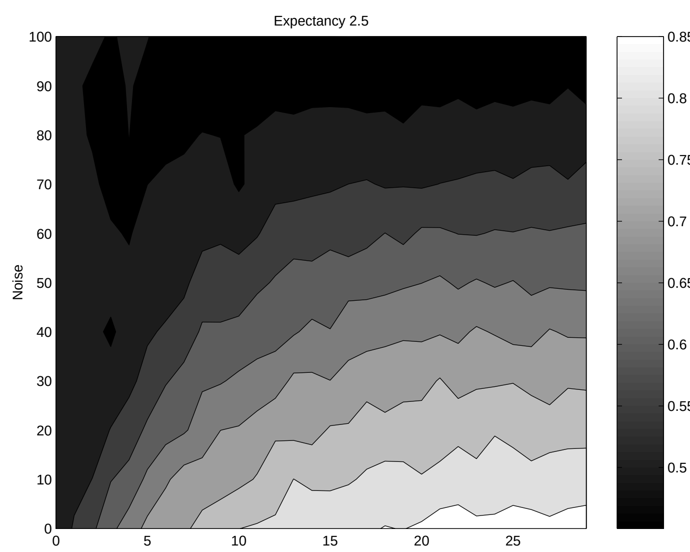

Fig. 3. Average density achieved by the algorithm for all resource constraints and all noise levels for an expectancy parameter of 2.5.

in the sequence may be erroneously executed. Also the sequence of moves can be stopped before completion, with increasing probability depending on the number of moves in the sequence.

In the simulations described next, we used a simplistic model of noise as the more involved model gives similar results.

D. Overview of the Experimental Conditions

We carried out simulations of the action selection mechanism for the packing task under constraints. We varied each of the following parameters:

- Expectancy was varied from a value of 0 to 5.0 in 0.1 increments. The chosen expectancy remained fixed for a 10 block trial.
- Resources were varied from 0 to 29. This represents the number of resources that can be expended in total for a 10 block trial. For example, 15 resources means that only a total of 15 moves and rotations can be performed for the 10 blocks in a trial.
- Noise was varied from 0% to 100% in 10% increments. As previously stated, the results presented here were obtained using a simple model of noise. Noise is intended to model the performance of subjects under increasing time constraints.

For each of the 3 combinations of parameters, 100 trials were run and the average performance on the 100 trials was calculated. Performance was rated by the density of the packing achieved.

E. Results of Computer Experiments

Figure 2 shows the results of the simulation for 4 values of noise: 0, 20, 40 and 60%. The 4 contour plots display the density achieved by the algorithm for all combinations of resource constraints and expectancy at a given noise level. As noise increases the level of performance decreases over all values of resources. Also, and not surprisingly, better performance is achieved under conditions of more resource availability.

The most striking feature of the results are the prominent difference that the expectancy parameter can make in performance. In particular there is a great increase in performance around an expectancy of 2.5, which is most prominent at 0% noise but is still visible at 20 and 40%. An expectancy parameter of 2.5 represents an optimal intuitive heuristic for decision making in the packing task as we have set it up. At 2.5 expectancy the algorithm achieves a good balance between conserving resources under conditions of tight constraints and choosing good moves when possible. Expectancy values above 3 still work, however they tend to be too conservative and performance begins to degrade. When the algorithm is too conservative it ends up with unused resources at the end of the trial.

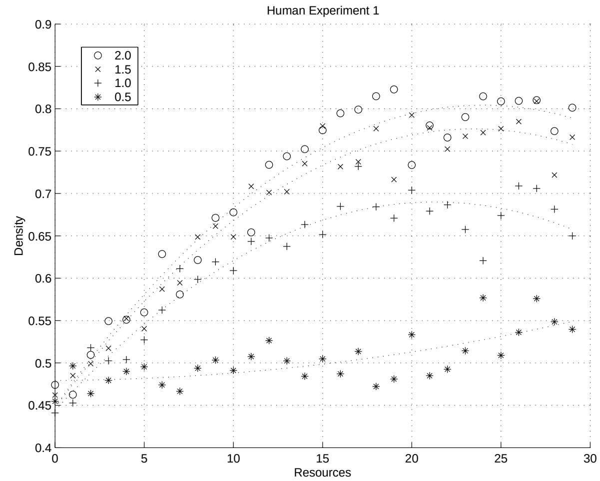

Fig. 4. Human performance on the packing task for the first set of trials. Time constraint conditions ranged from 2.0 to 0.5 seconds in 0.5 second intervals.

In figure 3 we show a different view of the results for a single value of expectancy. Figure 3 shows the performance over all values of noise and resources for an expectancy value of 2.5. This figure reveals that resource limitations above 20 have little effect on performance. But below 20, the achievable performance begins to fall. This indicates that more than 20 resources are needed in the usual case in order to achieve good packing densities on the task.

IV. HUMAN TRIALS

A. Overview of the Experimental Conditions

To develop our models of the parameters that people may intuitively learn and adapt when performing the packing task, we performed a series of packing trials on human subjects. Subjects were asked to perform many packing tasks, with differing time and resource constraints. Subjects were first allowed to practice on the task until they were comfortable that they had obtained a certain level of competence.

In the first set of experiments we ran 14 subjects. Each subject performed 30 packing trials with a 2.0 second time constraint, then 30 more with a 1.5 second time constraint, and similarly for 1.0 and 0.5 second time constraints. The time constraint set a limit on how much time they had to complete moving a block to its intended position before it was dropped for them. Each of the 30 trials for a particular time level consisted of performing a 10 block packing task at a different resource constraint level, which varied from 0 resources to 29 resources. The order that they received the resource constraint trials was varied randomly. So they might first perform a trial

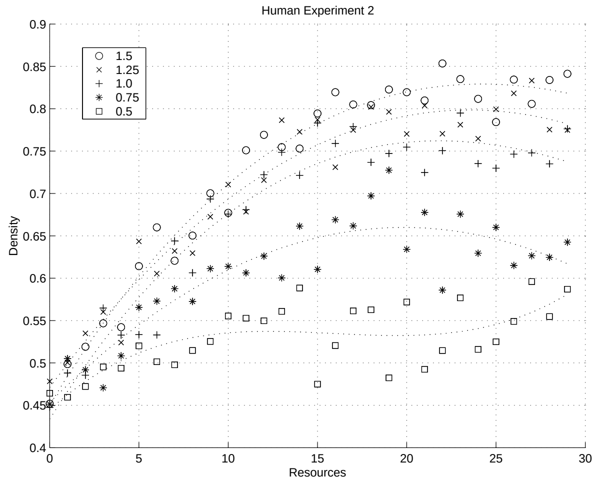

Fig. 5. Human performance in the second set of trials using 3 expert subjects. Time constraints ranged from 1.5 to 0.5 seconds in 0.25 second intervals.

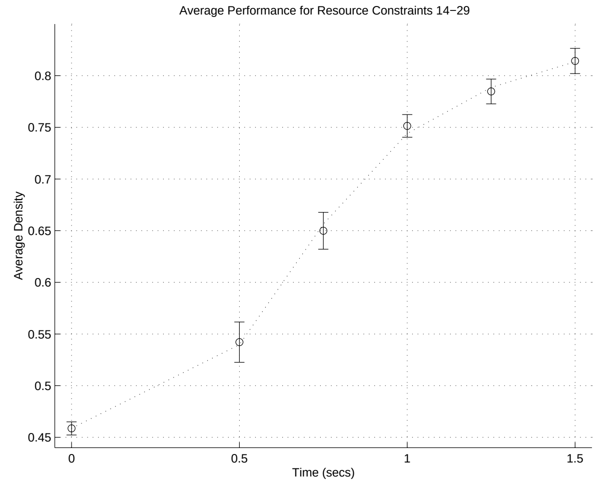

Fig. 6. Average performance achieved by expert human subjects over resource constraint conditions ranging from 14 to 29 resources. Notice the sharp drop in performance between 1.0 and 0.5 seconds.

with 15 resources, then with 5 resources, etc. 0 resource trials acted as a type of control that allowed for us to develop a minimum baseline density that happens on the packing task when blocks simply fall at random onto the playing field.

The second set of experiments had a similar set up as the first one. In the second set we ran 16 experiments using 3 subjects. The time constraints were varied from 1.5 to 0.5 seconds in 0.25 second intervals (1.5, 1.25, 1.0, 0.75, 0.5). Each time level had 30 trials with a different resource constraint as in the previous experiment. The 3 subjects who performed the 16 experiments were players who had achieved a high level of proficiency on the packing task, obtaining better performance on average than other players. The subjects reached this level of performance through repeated practice and experience with the task.

B. Results of Human Trials

Figure 4 displays the results of the human trials for the first experiment. This experiment was performed mainly to determine the critical time constraint range where humans are no longer able to sustain performance because the task is happening too fast for them to process. As shown in the figure, this point appears to happen somewhere between 1.0 and 0.5 seconds. From our observations of the trials, 1.0 second still allowed people enough time to perform some rotations and execute their intended sequence of actions. However at 0.5 seconds blocks fell so fast that the subjects could only react minimally, usually by trying to guide the blocks to the left or right with no attempts or possibility of performing rotations.

Figure 5 shows the results of the second human trials. We performed more trials in the critical region between 1.0 and 0.5 seconds. In the region between 1.0 and 0.5 seconds the human subject is sometimes able to complete their intended moves, but not always. Also error rates increase drastically in this region. Notice that the 3 expert level subjects performed better on average over the subjects in the first experiment. In fact, the subjects in this experiment have gained enough proficiency to push the critical time constraint level to a smaller value. In experiment 1 there appears to be a small drop in performance between 1.5 and 1.0 seconds. The expert subjects performed noticeably better at the 1.0 second time constraint level. They managed to push the critical time constraint down to somewhere at or below the 0.75 second constraint.

Figure 6 displays the average density achieved by the expert human subjects in the second experiment for resource constraint levels (from 14 to 29 resources) at each of the 5 time constraints. This figure illustrates the transition in performance that happens as the time constraint reaches a critical level beyond that of human brains to cope with. The data point at 0 seconds is the density that is achieved when blocks are randomly dropped onto the playing field (e.g. the player has no time whatsoever to try and position the block).

V. DISCUSSION

Viewing action selection as the self-organization of a goal attractor landscape has several implications. The timing of the dynamics to find and settle into an attractor basin sets limits on the real-time performance possible by biological brains. However, learning and experience can serve to deepen some attractors, with the effect that good goals and intentions are found and settled into more quickly. Experience on the task allows for subjects to recognize and settle into good behavior

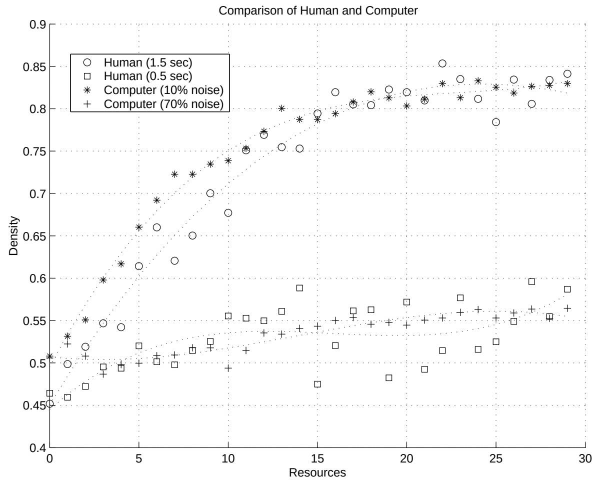

Fig. 7. Human performance from the second human trials for 1.5 and 0.5 seconds. We compare this to the computer simulation of action selection at 10% and 70% noise levels.

attractors more quickly, and therefore push back the threshold of time constraints under which they can produce effective behavior. Our expert subjects show some evidence of this ability to extend good performance into increasingly difficult time constraint domains.

The time to recognize and fall into a goal attractor is a critical feature of performing tasks under real time constraints. In the recognition of perceptual categories, two types of emergent amplitude modulation (AM) patterns have been identified [10], [11], [12], [4]. When a stimulus is given to a system, there is a phase transition from a high dimensional chaotic attractor to a lower dimensional wing. The first type of AM pattern occurs with a short latency immediately after the stimulus arrives. This early AM patterns represents the impact of a discriminated stimulus on the activity of the receiving cortex. The second type of AM pattern is endogenous (e.g. internally generated) and occurs with a variable latency in the time range of between 750 and 1200 ms.

The second type of emergent pattern represents the act of categorization of the input stimulus. The type II pattern is the result of recognition of stimuli meaningful in the current situation to the organism.

We believe that the same type of pattern formation works in the emergence of intentional actions from the entorhinal cortex. In this case, the dynamics provide the action selection mechanism for recognizing and choosing among strategies and goals for the organism. The time period of around 750 ms. represents the minimum time that the selection and performance of intentional goal actions can be achieved. Through learning in real time tasks, people can push their performance up to this limit, but the fundamental dynamics of the brain dictates that intentional goals cannot be formed in time frames less than this period needed for the formation of type II AM patterns.

Finally we speculate on the possible connection between our computational models and human performance under time constraints. Figure 7 indicates that there is a reasonable match between computer simulations with 10% noise and human performance with small time pressures (1.5 sec available). On the other hand, 70% noise produces performance close to humans with very significant time pressures (0.5 sec).

ACKNOWLEDGEMENTS

This work was supported by NASA Intelligent Systems Research Grant NCC-2-1244. Discussions with Dr. Stan Franklin and Dr. Phillip Wolff were greatly helpful and appreciated. We would also like to acknowledge the student volunteers of COMP/PSYCH 7/8514 for their help as subjects.

References

[1] Stanley P. Franklin. *Artificial Minds.* The MIT Press, Cambridge, MA, 1995.

[2] J. A. Scott Kelso. *Dynamic Patterns: The Self-organization of Brain and Behavior.* The MIT Press, Cambridge, MA, 1995.

[3] Christine A. Skarda and Walter J. Freeman. How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, 10:161–195, 1987.

[4] Robert Kozma and Walter J. Freeman. Encoding and recall of noisy data as chaotic spatio-temporal memory patterns in the style of the brains. In *Proceedings of the IEEE/INNS International Joint Conference on Neural Networks*, pages 52–57, 2000.

[5] Esther Thelen and Linda B. Smith. *A Dynamic Systems Approach to the Development of Cognition and Action.* The MIT Press, Cambridge, MA, 1994.

[6] Walter J. Freeman. *How Brains Make Up Their Minds.* Weidenfeld & Nicolson, London, 1999.

[7] Derek Harter and Robert Kozma. Models of ontogenetic development for autonomous adaptive systems. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, Edinburgh, Scotland, August 2001.

[8] Derek Harter and Robert Kozma. Task environments for the dynamic development of behavior. In *Proceedings of the Intelligent Systems Design and Applications 2001 Workshop (ISDA 2001)*, San Francisco, CA, May 2001.

[9] David Kirsh and Paul Maglio. Reaction and reflection in tetris. In J. Hendler, editor, *Artificial Intelligence Planning Systems: Proceedings of the First Annual International Conference (AIPS92)*, San Mateo, CA, 1992. Morgan Kaufman.

[10] Robert Kozma and Walter J. Freeman. Classification of EEG patterns using nonlinear dynamics and identifying chaotic phase transitions. *CNS*2001 Special Issue of Neurocomputing*, 2001.

[11] Robert Kozma and Walter J. Freeman. Chaotic resonance - methods and applications for robust classification of noisy and variable patterns. *International Journal of Bifurcation and Chaos*, 11(6), 2001.

[12] Robert Kozma and Walter J. Freeman. A possible mechanism for intermittent oscillations in the KIII model of dynamic memories - the case study of olfaction. In *Proceedings IJCNN 1999*, pages 52–57, 1999.