Intelligence through interaction: Towards a unified theory for learning

Ah-hwee TAN
Singapore Management University, ahtan@smu.edu.sg

Gail A. CARPENTER

Stephen GROSSBERG

Intelligence Through Interaction: Towards a Unified Theory for Learning

Ah-Hwee Tan[1], Gail A. Carpenter[2], and Stephen Grossberg[2]

[1] Intelligent Systems Centre and School of Computer Engineering
Nanyang Technological University, Nanyang Avenue, Singapore 639798
asahtan@ntu.edu.sg

[2] Center for Adaptive Systems and Department of Cognitive and Neural Systems
Boston University, 677 Beacon Street, Boston, MA 02215, USA
gail,steve@cns.bu.edu

Abstract. Machine learning, a cornerstone of intelligent systems, has typically been studied in the context of specific tasks, including clustering (unsupervised learning), classification (supervised learning), and control (reinforcement learning). This paper presents a learning architecture within which a universal adaptation mechanism unifies a rich set of traditionally distinct learning paradigms, including learning by matching, learning by association, learning by instruction, and learning by reinforcement. In accordance with the notion of embodied intelligence, such a learning theory provides a computational account of how an autonomous agent may acquire the knowledge of its environment in a real-time, incremental, and continuous manner. Through a case study on a minefield navigation domain, we illustrate the efficacy of the proposed model, the learning paradigms encompassed, and the various types of knowledge learned.

1 Introduction

Machine learning, a cornerstone of intelligent system research, has typically been studied in the context of specific tasks, including clustering (unsupervised learning), classification (supervised learning), and control (reinforcement learning). In reality, an autonomous system acquires intelligence through its interaction with the environment. This is in keeping with the view in modern cognitive science that cognition is a process deeply rooted in the body's interaction with the world [1]. Embodied cognition is also akin to the intensive study on reinforcement learning [15] in which an autonomous agent learns to adjust its behaviour according to evaluative feedback received from the environment.

Over the past decades, a family of neural architectures known as Adaptive Resonance Theory (ART) [3,5,8,9], has been steadily developed. With well-founded computational principles, ART has been applied successfully to many pattern analysis, recognition, and prediction applications [6,12]. These successful applications are of particular interest because the basic ART principles have been derived from an analysis of human and animal perceptual and cognitive

Published in Advances in Neural Networks: International Symposium on Neural Networks (ISNN) 2007: Nanjing, June 3-7, Proceedings. Springer, LNCS 4492, pp. 1094-1103.
DOI: 10.1007/978-3-540-72383-7_128

information processing, and have led to behavioral and neurobiological predictions that have received significant experimental support during the last decade; see Grossberg 2003 and Raizada & Grossberg 2003 for reviews. In this paper, we show that Adaptive Resonance Theory lays the foundation of a unified model that encompasses a myriad of learning paradigms, traditionally viewed as distinct. The proposed model is a natural extension of the original ART models from a single pattern field to multiple pattern channels. Whereas the original ART models [2] perform unsupervised learning of recognition nodes in response to incoming input patterns, the proposed neural architecture, known as fusion ART (fusion Adaptive Resonance Theory), learns multi-channel mappings simultaneously across multi-modal pattern channels in an online and incremental manner.

To illustrate the unified model, this paper presents a case study based on a minefield navigation task, which involves an autonomous vehicle (AV) learning to navigate through obstacles to reach a stationary target (goal) within a specified number of steps. The experimental results show that fusion ART is capable of performing a myriad of learning tasks and is able to produce a fast and stable learning performance.

The rest of the paper is organized as follows. Section 2 provides a summary of the fusion ART architecture and the associated system dynamics. Sections 3, 4, 5 and 6 show how fusion ART can be used for various types of learning tasks. Section 7 illustrates the fusion ART functionalities and performance based on the minefield navigation task. The final section concludes and highlights possible future directions.

2 Fusion ART

Fusion ART employs a multi-channel architecture (Figure. 1), comprising a category field F_2 connected to a fixed number of (K) pattern channels or input fields through bidirectional conditionable pathways. The model unifies a number of network designs, most notably Adaptive Resonance Theory (ART) [3,5], Adaptive Resonance Associative Map (ARAM) [16] and Fusion Architecture for Learning, COgnition, and Navigation (FALCON) [20], developed over the past decades for a wide range of functions and applications. The generic network dynamics of fusion ART, based on fuzzy ART operations [4], is summarized as follows.

Input vectors: Let $\mathbf{I}^{ck} = (I_1^{ck}, I_2^{ck}, \ldots, I_n^{ck})$ denote the input vector, where $I_i^{ck} \in [0,1]$ indicates the input i to channel ck. With complement coding, the input vector $\mathbf{I}^{ck}$ is augmented with a complement vector $\bar{\mathbf{I}}^{ck}$ such that $\bar{I}_i^{ck} = 1 - I_i^{ck}$. **Activity vectors:** Let $\mathbf{x}^{ck}$ denote the F_1^{ck} activity vector for $k = 1, \ldots, K$. Let $\mathbf{y}$ denote the F_2 activity vector. **Weight vectors:** Let $\mathbf{w}_j^{ck}$ denote the weight vector associated with the jth node in F_2 for learning the input patterns in F_1^{ck} for $k = 1, \ldots, K$. Initially, F_2 contains only one *uncommitted* node and its weight vectors contain all 1's. **Parameters:** The fusion ART's dynamics is determined by choice parameters $\alpha^{ck} > 0$, learning rate parameters $\beta^{ck} \in [0,1]$, contribution parameters $\gamma^{ck} \in [0,1]$ and vigilance parameters $\rho^{ck} \in [0,1]$ for $k = 1, \ldots, K$.

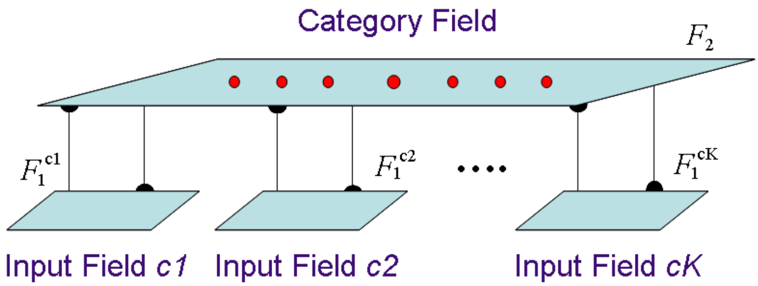


Fig. 1. The fusion ART architecture

As a natural extension of ART, fusion ART responds to incoming patterns in a continuous manner. It is important to note that at any point in time, fusion ART does not require input to be present in all the pattern channels. For those channels not receiving input, the input vectors are initialized to all 1s. The fusion ART pattern processing cycle comprises five key stages, namely code activation, code competition, activity readout, template matching, and template learning, as described below. **Code activation:** Given the activity vectors $\mathbf{I}^{c1}, \ldots, \mathbf{I}^{cK}$, for each F_2 node j, the choice function T_j is computed as follows:

$$T_j = \sum_{k=1}^{K} \gamma^{ck} \frac{|\mathbf{I}^{ck} \wedge \mathbf{w}_j^{ck}|}{\alpha^{ck} + |\mathbf{w}_j^{ck}|}, \tag{1}$$

where the fuzzy AND operation $\wedge$ is defined by $(\mathbf{p} \wedge \mathbf{q})_i \equiv min(p_i, q_i)$, and the norm $|.|$ is defined by $|\mathbf{p}| \equiv \sum_i p_i$ for vectors $\mathbf{p}$ and $\mathbf{q}$. **Code competition:** A code competition process follows under which the F_2 node with the highest choice function value is identified. The winner is indexed at J where

$$T_J = \max\{T_j : \text{for all } F_2 \text{ node } j\}. \tag{2}$$

When a category choice is made at node J, $y_J = 1$; and $y_j = 0$ for all $j \neq J$. This indicates a winner-take-all strategy. **Activity readout:** The chosen F_2 node J performs a readout of its weight vectors to the input fields F_1^{ck} such that

$$\mathbf{x}^{ck} = \mathbf{I}^{ck} \wedge \mathbf{w}_J^{ck}. \tag{3}$$

Template matching: Before the activity readout is stabilized and node J can be used for learning, a template matching process checks that the weight templates of node J are sufficiently close to their respective input patterns. Specifically, resonance occurs if for each channel k, the *match function* m_J^{ck} of the chosen node J meets its vigilance criterion:

$$m_J^{ck} = \frac{|\mathbf{I}^{ck} \wedge \mathbf{w}_J^{ck}|}{|\mathbf{I}^{ck}|} \geq \rho^{ck}. \tag{4}$$

If any of the vigilance constraints is violated, mismatch reset occurs in which the value of the choice function T_J is set to 0 for the duration of the input presentation. Using a *match tracking* process, at the beginning of each input presentation, the vigilance parameter ρ^{ck} in each channel ck equals a baseline vigilance $\bar{\rho}^{ck}$. When a mismatch reset occurs, the ρ^{ck} of all pattern channels are increased simultaneously until one of them is slightly larger than its corresponding match function m_J^{ck}, causing a reset. The search process then selects another F_2 node J under the revised vigilance criterion until a resonance is achieved. **Template learning:** Once a resonance occurs, for each channel ck, the weight vector $\mathbf{w}_J^{ck}$ is modified by the following learning rule:

$$\mathbf{w}_J^{ck(\text{new})} = (1 - \beta^{ck})\mathbf{w}_J^{ck(\text{old})} + \beta^{ck}(\mathbf{I}^{ck} \wedge \mathbf{w}_J^{ck(\text{old})}). \tag{5}$$

When an uncommitted node is selected for learning, it becomes *committed* and a new uncommitted node is added to the F_2 field. Fusion ART thus expands its network architecture dynamically in response to the input patterns.

The network dynamics described above can be used to support a myriad of learning operations. We show how fusion ART can be used for a variety of traditionally distinct learning tasks in the subsequent sections.

3 Learning by Similarity Matching

With a single pattern channel, the fusion ART architecture reduces to the original ART model. Using a selected vigilance value ρ, an ART model learns a set of recognition nodes in response to an incoming stream of input patterns in a continuous manner. Each recognition node in the F_2 field learns to encode a template pattern representing the key characteristics of a set of patterns. ART has been widely used in the context of unsupervised learning for discovering pattern groupings. Please refer to the selected ART literatures [3,5,8,9] for a review of ART's functionalities, interpretations, and applications.

4 Learning by Association

By synchronizing pattern coding across multiple pattern channels, fusion ART learns to encode associative mappings across distinct pattern spaces. A specific instance of fusion ART with two pattern channels is known as Adaptive Resonance Associative Map (ARAM), that learns multi-dimensional supervised mappings from one pattern space to another pattern space [16]. An ARAM system consists of an input field F_1^a, an output field F_1^b, and a category field F_2. Given a set of feature vectors presented at F_1^a with their corresponding class vectors presented at F_1^b, ARAM learns a predictive model (encoded by the recognition nodes in F_2) that associates combinations of key features to their respective classes.

Fuzzy ARAM, based on fuzzy ART operations, has been successfully applied to numerous machine learning tasks, including personal profiling [19], document classification [11], personalized content management [18], and DNA gene expression analysis [22]. In many benchmark experiments, ARAM has demonstrated predictive performance superior to those of many state-of-the-art machine learning systems, including C4.5, Backpropagation Neural Network, K Nearest Neighbour, and Support Vector Machines.

5 Learning by Instruction

During learning, fusion ART formulates recognition categories of input patterns across multiple channels. The knowledge that fusion ART discovers during learning, is compatible with symbolic rule-based representation. Specifically, the recognition categories learned by the F_2 category nodes are compatible with a class of IF-THEN rules that maps a set of input attributes (antecedents) in one pattern channel to a disjoint set of output attributes (consequents) in another channel. Due to this compatibility, at any point of the incremental learning process, instructions in the form of IF-THEN rules can be readily translated into the recognition categories of a fusion ART system. The rules are conjunctive in the sense that the attributes in the IF clause and in the THEN clause have an *AND* relationship. Augmenting a fusion ART network with domain knowledge through explicit instructions serves to improve learning efficiency and predictive accuracy.

The fusion ART rule insertion strategy is similar to that used in Cascade ARTMAP, a generalization of ARTMAP that performs domain knowledge insertion, refinement, and extraction [17]. For direct knowledge insertion, the IF and THEN clauses of each instruction (rule) is translated into a pair of vectors $\mathbf{A}$ and $\mathbf{B}$ respectively. The vector pairs derived are then used as training patterns for inserting into a fusion ART network. During rule insertion, the vigilance parameters are set to 1s to ensure that each distinct rule is encoded by one category node.

6 Learning by Reinforcement

Reinforcement learning [15] is a paradigm wherein an autonomous system learns to adjust its behaviour based on reinforcement signals received from the environment. An instance of fusion ART, known as FALCON (Fusion Architecture for Learning, COgnition, and Navigation), learns mappings simultaneously across multi-modal input patterns, involving states, actions, and rewards, in an online and incremental manner. Compared with other ART-based reinforcement learning systems, FALCON presents a truly integrated solution in the sense that there is no implementation of a separate reinforcement learning module or Q-value table. Using competitive coding as the underlying principle of computation, the network dynamics encompasses a myriad of learning paradigms, including unsupervised learning, supervised learning, as well as reinforcement learning.

FALCON employs a three-channel architecture, comprising a category field F_2 and three pattern fields, namely a sensory field F_1^{c1} for representing current states, a motor field F_1^{c2} for representing actions, and a feedback field F_1^{c3} for representing reward values. A class of FALCON networks, known as TD-FALCON [21,23], incorporates Temporal Difference (TD) methods to estimate and learn value function $Q(s,a)$, that indicates the goodness to take a certain action a in a given state s.

The general sense-act-learn algorithm for TD-FALCON is summarized in Table 1. Given the current state s, the FALCON network is used to predict the value of performing each available action a in the action set $\mathcal{A}$ based on the corresponding state vector **S** and action vector **A**. The value functions are then processed by an action selection strategy (also known as policy) to select an action. Upon receiving a feedback (if any) from the environment after performing the action, a TD formula is used to compute a new estimate of the Q-value for performing the chosen action in the current state. The new Q-value is then used as the teaching signal (represented as reward vector **R**) for FALCON to learn the association of the current state and the chosen action to the estimated value.

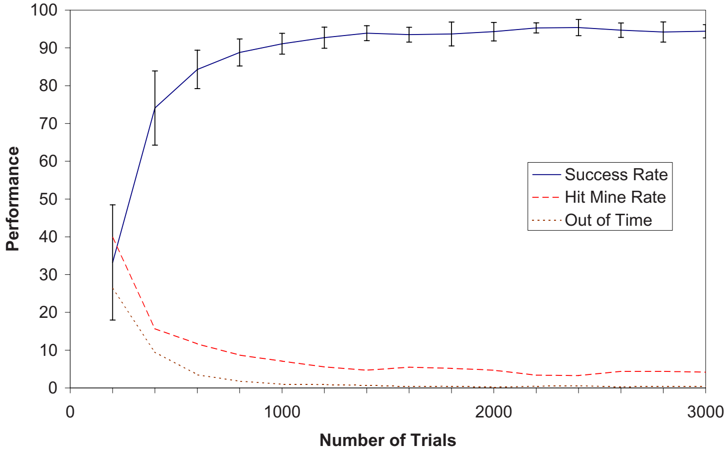
Table 1. The TD−FALCON algorithm

1. Initialize the FALCON network.
2. Given the current state s, for each available action a in the action set $\mathcal{A}$, predict the value of performing the action Q(s,a) by presenting the corresponding state and action vectors **S** and **A** to FALCON.
3. Based on the value functions computed, select an action a from $\mathcal{A}$ following an action selection policy.
4. Perform the action a, observe the next state s', and receive a reward r (if any).
5. Estimate the value function $Q(s,a)$ following a temporal difference formula given by $\Delta Q(s,a) = \alpha TD_{err}$.
6. Present the corresponding state, action, and reward (Q-value) vectors, namely **S**, **A**, and **R**, to FALCON for learning.
7. Update the current state by s=s'.
8. Repeat from Step 2 until s is a terminal state.

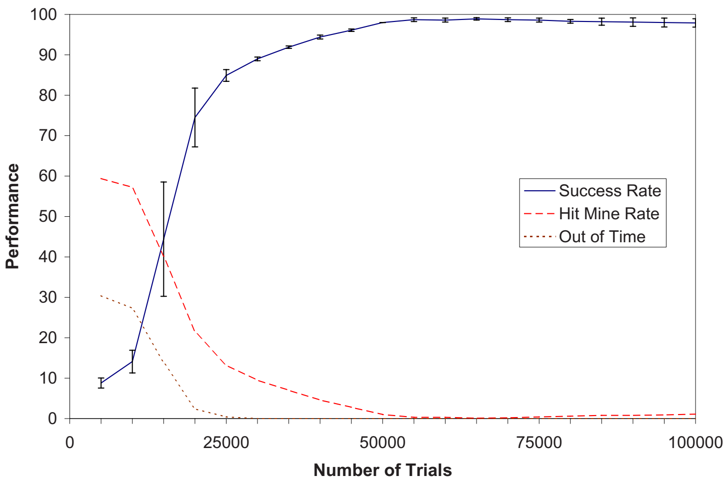
7 Case Study: Minefield Navigation

The minefield simulation task studied here is similar to the underwater navigation and mine avoidance domain developed by Naval Research Lab (NRL) [7,14]. The objective is to navigate through a minefield to a randomly selected target position in a specified time frame without hitting a mine. In each trial, the autonomous vehicle (AV) starts from a random position in the field, and repeats the cycles of sense, act, and learn. A trial ends when the system reaches the target (success), hits a mine (failure), or exceeds 30 sense-act-learn cycles (out of time). The target and the mines remain stationary during the trial.

Minefield navigation and mine avoidance is a non-trivial task. As the configuration of the minefield is generated randomly and changes over trials, the



(a) TD-FALCON



(b) BP-Q Learner

Fig. 2. The task completion performance of TD-FALCON compared with the BP-Q Learner operating with delayed rewards

system needs to learn strategies that can be carried over across experiments. For sensing, the AV has a coarse 180 degree forward view based on five sonar sensors. In each direction i, the sonar signal is measured by $s_i = \frac{1}{d_i}$, where d_i is the distance to an obstacle in the i direction. Other sensory inputs include the current and target bearings. In each step, the system chooses one of the five possible actions, namely move left, move diagonally left, move straight ahead, move diagonally right, and move right.

In this domain, we conduct experiments using TD-FALCON, a three-channel fusion ART model, with both immediate and delayed evaluative feedback. For both reward schemes, at the end of a trial, a reward of 1 is given when the AV reaches the target. A reward of 0 is given when the AV hits a mine. For the immediate reward scheme, a reward is estimated at each step of the trial by computing a utility function $utility = \frac{1}{1+rd}$, where rd is the remaining distance between the AV and the target position.

7.1 Performance Comparison

We compare the performance of TD-FALCON with an alternative reinforcement learning system (hereafter referred to as the BP-Q Learner), in terms of success rate, hit-mine rate, and out-of-time rate in a 16 by 16 minefield containing 10 mines. The BP-Q Learner uses the standard Q-learning rule and a gradient descent based multi-layer feedforward neural network as the function approximator. For illustration purpose, we only show the performance of the two systems operating with delayed rewards (Figure. 2). We can see that the BP-Q Learner generally takes a very large number of (more than $40,000$) trials to reach 90% success rates. In contrast, TD-FALCON consistently achieves the same level of performance within the first 1000 trials. In other words, TD-FALCON learns at least an order of magnitude faster than the BP-Q learner.

Considering network complexity, the BP-Q Learner has the advantage of a highly compact network architecture. When trained properly, a BP network consisting of 36 hidden nodes can produce performance equivalent to that of a TD-FALCON model with say 200 category nodes. In terms of the speed of adaptation, however, TD-FALCON is clearly a faster learner by consistently mastering the task in a much smaller number of trials.

7.2 Knowledge Interpretation

To illustrate the variety of the knowledge learned by TD-FALCON, Table 2 shows a sample set of the knowledge encoded by its recognition nodes. Through learning by similarity matching, TD-FALCON identifies key situations in its environment that are of significance to its mission. Two such typical situations are shown in the first row of the table. Through learning by association (or directly as instructions), TD-FALCON learns the association between typical situations and their corresponding desired actions. Two such association rules are shown in the second row. Finally, through the reinforcement signals given by the environment, TD-FALCON learns the value of performing a specific action in a given situation. The third row shows two extreme cases, one indicating a high payoff for taking an action in a situation and the other giving a severe penalty for taking the same action in a slightly different situation.

Table 2. Sample knowledge learned by FALCON in the minefield navigation domain. $\wedge$ is used here to indicate *AND* operator.

Type of Learning	Knowledge Learned
Similarity matching	FrontSonar=1.0 $\wedge$ Target=Front FrontSonar$\leq$0.5 $\wedge$ Target=Front
Association or Instruction	**IF** FrontSonar$\leq$0.5 $\wedge$ Target=Front **THEN** Move=Front **IF** FrontSonar=1.0 $\wedge$ DRightSonar$\leq$0.5 $\wedge$ Target=Front **THEN** Move=DRight
Reinforcement	**IF** FrontSonar$\leq$0.5 $\wedge$ Target=Front **THEN** Move=Front (Q=1.0) **IF** FrontSonar=1.0 $\wedge$ Target=Front **THEN** Move=Front (Q=0.0)

8 Conclusion

This paper has outlined a generalized neural architecture, known as fusion Adaptive Resonance Theory (fusion ART), that learns multi-dimensional mappings simultaneously across multi-modal pattern channels, in an online and incremental manner. Such a learning architecture enables an autonomous agent to acquire its intelligence in a real-time dynamic environment. Using Adaptive Resonance Theory (ART) as an universal coding mechanism, the proposed model unifies a myriad of traditionally distinct learning paradigms, including unsupervised learning, supervised learning, rule-based knowledge integration, and reinforcement learning. In fact, ART-style learning and matching mechanism seems to be operative in many levels of the cerebral cortex of the brain, especially in the vision system [10]. The proposed framework may thus serve as a foundation model for developing high level cognitive information processing capabilities, including awareness, reasoning, explaining, and surprise handling.

References

1. Anderson, M.L.: Embodied Cognition: A Field Guide. Artificial Intelligence **149** (2003) 91-130
2. Carpenter, G. A., Grossberg, S.: A Massively Parallel Architecture for a Self- organizing Neural Pattern Recognition Machine. Computer Vision, Graphics, and Image Processing **37** (1987) 54-115
3. Carpenter, G. A., Grossberg, S., editors: Pattern Recognition by Self-Organizing Neural Networks. Cambridge, MA: MIT Press (1991)
4. Carpenter, G. A., Grossberg, S., Rosen, D. B.: Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System. Neural Networks **4** (1991) 759-771
5. Carpenter, G.A., Grossberg, S.: Adaptive Rresonance Theory. In M.A. Ar- bib, editor, The Handbook of Brain Theory and Neural Networks, 87-90. Cambridge, MA: MIT Press (2003)
6. Duda, R. O., Hart, P.E., Stock, D.G., editors. Pattern Classification (2nd edi- tion). New York: John Wiley, Section 10.11.2, (2001)

7. Gordan, D., Subramanian, D.: A Cognitive Model of Learning to Navigate. In Proceedings, Nineteenth Annual Conference of the Cognitive Science Society (1997) 271-276
8. Grossberg, S.: Adaptive Pattern Recognition and Universal Recoding, I: Parallel Development and Coding of Neural Feature Detectors. Biological Cybernetics **23** (1976) 121-134
9. Grossberg, S.: Adaptive Pattern Recognition and Universal Recoding, II: Feedback, Expectation, Olfaction, and Illusion. Biological Cybernetics **23** (1976) 187-202
10. Grossberg, S.: How Does the Cerebral Cortex Work? Development, Learning, Attention, and 3d Vision by Laminar Circuits of Visual cortex. Behavioral and Cognitive Neuroscience Reviews **2** (2003) 47-76
11. He, J., Tan, A-H., Tan, C-L: On Machine Learning Methods for Chinese Documents Classification. Applied Intelligence: Special Issue on Text and Web Mining **18** (3) (2003) 311-322
12. Levine, D. S., editor: Introduction to Neural and Cognitive Modeling. New Jersey: Lawrence Erlbaum Associates, Chapter 6 (2000)
13. Raizada, R., Grossberg, S.: Towards a Theory of the Laminar Architecture of Cerebral Cortes: Computational Clues from the Visual System. Cerebral Cortex **13** (2003) 200-213
14. Sun, R., Merrill, E., Peterson, T.: From Implicit Skills to Explicit Knowledge: A Bottom-up Model of Skill Learning. Cognitive Science **25** (2) (2001) 203-244
15. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press (1998)
16. Tan, A. H.: Adaptive Resonance Associative Map. Neural Networks **8** (3) (1995) 437-446
17. Tan, A. H.: Cascade ARTMAP: Integrating Neural Computation and Symboli Knowledge Processing. IEEE Transactions on Neural Networks **8** (2) (1997) 237-250
18. Tan, A-H., Ong, H-L., Pan, H., Ng, J., Li, Q-X.: Towards Personalized Web Intelligence. Knowledge and Information Systems **6** (5) 595-616 (2004)
19. Tan, A-H., Soon, H.-S.: Predictive Adaptive Resonance Theory and Knowledge Discovery in Database. In Proceedings, Fourth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD'00), Kyoto, pages (2000) 173-176
20. Tan, A.H.: FALCON: A Fusion Architecture for Learning, Cognition, and Navigation. In Proceedings, International Joint Conference on Neural Networks (2004) 3297-3302
21. Tan, A.H.: Self-organizing Neural Architecture for Reinforcement Learning. In Proceedings, J. Wang et al. (Eds.): International Symposium on Neural Networks (ISNN'06), Chengdu, China, LNCS **3971** (2006) 470-475
22. Tan, A.H., Pan, H.: Predictive Neural Networks for Gene Expression Data Analysis. Neural Networks **18** (3) (2005) 297-306
23. Tan, A.H., Lu, N., Xiao, D.: Integrating Temporal Difference Methods and Self-organizing Neural Networks for Reinforcement Learning with Delayed Evaluative feedback. IEEE Transactions on Neural Networks, to appear