Teaching an Agent by Playing a Multimodal Memory Game: Challenges for Machine Learners and Human Teachers

Byoung-Tak Zhang

Biointelligence Laboratory
School of Computer Science and Engineering &
Graduate Programs in Cognitive Science and Brain Science
Seoul National University, Seoul 151-742, Korea
http://bi.snu.ac.kr/~btzhang/

Abstract

As agents become ubiquitous in virtual as well as physical worlds, the importance of learning from real-life human interaction is increasing. Here we explore new learning and teaching strategies for an agent situated in a digital cinema environment to solve a language-vision translation problem by playing a multimodal memory game with humans. We discuss the challenges for machine learners, i.e. learning architectures and algorithms, required to deal with this kind of long-lasting, dynamic scenario. We also discuss the challenges for human teachers to address the new machine learning issues. Based on our preliminary experimental results using the hypernetwork learning architecture we argue for self-teaching cognitive agents that actively interact with humans to generate queries and examples to evaluate and teach themselves.

Introduction

Most of existing machine learning algorithms assume a static environment where a fixed set of training examples are given, and the learning task is defined as optimization of some objective function, such as likelihood or entropy. However, this assumption is violated in many real world situations, where agents interact with humans in everyday life (Gibbs, 2005). The target function is generally *dynamic*, not static, since the humans may provide different feedbacks for similar situations at different times. The learning algorithm should be able to deal with *moving targets* since the target concept may drift. The interaction is typically a *long-lasting* process. The agent should be able to learn from *online* data. Also, the interaction is *multimodal* since multimodality is more natural and convenient for human teachers. The agents also need to learn to *integrate* multisensory data.

We have been developing a research platform that implements a cognitive game called multimodal memory game (MMG) to study machine learning architectures and algorithms for learning in a long-lasting, multimodal, interactive, and dynamic environment (Zhang, 2008a). The game is played by two or more human teachers and a machine learner. The participants watch movies or TV dramas and teach the machine to generate text from image and to produce image from text. In a simple version, the agent learns by observing the humans playing the game. In an advanced version, the learner actively participates in the game and learns by asking questions to the human players.

In this paper we use the MMG cognitive game platform as a think-pad for discussing the technical challenges and the division of roles between human teachers and machine learners. In the next section, we describe the MMG environment and discuss the technical issues to be solved by future machine learning technology. We then suggest three approaches to solving these problems, namely developing (i) new learning architectures, (ii) innovative learning strategies, and (iii) novel teaching strategies. We provide initial solutions to the first approach based on our previous work, i.e. using the hypernetwork learning model, to motivate further discussion on the approaches for (ii) and (iii).

This paper focuses on exploring alternative strategies for machine learners and human teachers in the context of the MMG game platform. Though this environment is limited to a digital cinema domain, the interaction between humans and machines, on one hand, and their interaction with multimodal media, on the other hand, can be made arbitrarily complex, and we believe this will shed light on understanding and designing teaching strategies for virtual and physical agents in a more general everyday life situation.

Multimodal Memory Game

The game consists of two humans and a machine learner situated in a digital cinema (**Fig. 1**). The goal of the machine is to learn the relationship of movie dialogues and pictures, i.e. the two modalities of language and vision, so that given an input query in one modality it produces the output in a different modality. For example, given a picture of the movie, the learner generates the sentences produced in the picture. Note that this problem is much harder than the sentence-picture matching task (MacLeod *et al*., 1978) used for IQ tests since the answer is not a verification of a relationship but a description of the relationship.

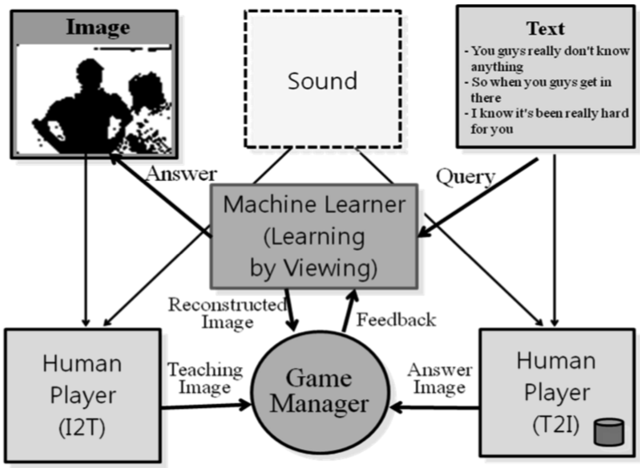

Fig. 1: The multimodal memory game (MMG). The participants, i.e. two human players and a machine agent, watch a movie together and play a memory game about the movie. In the text-to-image translation scenario described here, the human player T2I has to produce an image given a text query and the human player I2T evaluates this answer. The machine learner observes the play to learn to generate an image from a corresponding text. Thus, the agent learns by viewing the video and observing the human players engaging in the game.

All the players including the machine watch the movie or TV drama, such as "Frends". After watching, the humans play the game by question-and-answering about the movie pictures and dialogues. The task of one player, called I2T (for image-to-text), is to generate a text given a movie cut (image). The other player, named T2I (for text-to-image), is to generate an image given a text from the movie captions. While one player is asking, the other is answering. The two players alternate their roles. When the player is in the questioning mode, he receives all the multimodal inputs. When the player is in an answering mode, he receives the multimodal inputs except the modality in which he has to give an answer. The goal of the learner in this "multimodal memory game" is to imitate the human players by "learning by viewing", i.e. watching the movies, reading the captions, listening to the sounds, and observing the players engaging in the games over their shoulders.

The game can be played for an arbitrary long period by watching further movies. Also, the format of the game can be modified to make the task harder and to scale up the problem size, making a flexible research platform for investigating machine learning architectures and algorithms. In particular, the complexity of the tasks can be increased by extending the game in the following dimensions.

- New modality. Additional modalities can be added. For example, if a haptic sensor is available it might be incorporated into the memory model.
- More players. The number of players can be increased, making it a social learning game. This will enhance the diversity and reliability of the answers of the human teachers. Learning can be faster and more consistent.
- New media. The media contents can be extended. We can use the web UCC documents instead of the DVD videos. In the domain of education, the documents can be educational material for the school children.
- Gaming strategy. More variability can be added to increase the fun or responsibility of the participants in the games. For example, the gamers are allowed to watch the next scene only if they have passed the tests for the previous scenes.

By controlling the level of difficulty in the above dimensions, the platform can be tuned for long-term human-level performance studies as well as for short-term applications of practical interest.

Challenges for Machine Learners

The "learning by viewing" paradigm described above is a real challenge for current machine learning technology. First, the game is played *online* and the data are received in a *stream*. This is contrasted with typical supervised or unsupervised machine learning situations where a fixed set of training data is given. Second, the amount of data is *huge*. For some machine learning algorithms, such as support vector machines, there is a limit in the number of training examples. Third, the problem involves *multimodality*. It is necessary to integrate multiple sources of data into a coherent knowledge. Fourth, the game involves *vision and language* problems. Some teaching strategies are necessary to solve the problem without solving the harder symbol grounding problems (Chen & Mooney, 2008).

These issues can be addressed in several ways. Here we propose to approach them in three perspectives.

(i) Learning architecture (model)
(ii) Learning strategies (algorithms)
(iii) Teaching strategies (for humans)

The first approach tries to develop a new model architecture of the learner. The second approach is to develop new learning strategies for a given architecture. The third approach attempts to design intelligent teaching strategies for the human teachers. This is contrasted to the previous two approaches in that the solution is provided by the human teacher rather than by the machine learner.

The next section addresses the first approach. The following two sections discuss the second and third approaches. It should be mentioned already that these three approaches are not independent and can be combined to be more effective.

The Hypernetwork Learning Architecture

In this section we present an approach for addressing the new challenges based on our previous work. We demonstrate that a version of the MMG game can be solved by the hypernetwork learning model (Zhang, 2008b), and then discuss the remaining challenges to solve the problem in a more real-life setting.

The gameplay was simulated by logging a collection of approximately 3000 pairs of sentences and the corresponding pictures of TV drama “Friends”. The image data are preprocessed to generate the *visual words* or a visual vocabulary V_I of size 4800 (= 80 pixels by 60 pixels). The text data are converted to a linguistic vocabulary V_T of approximately 2500 words.

The general learning scheme is illustrated in **Fig. 2**. From a given example of text-image pair, a large number of random lists of words and visual words are sampled. For example, from a scene 1 consisting of “there’s nothing to tell” and “*v*1 *v*2 *v*3 v4”, where *v*i are visual words, a list “there’s nothing v1 v2 v4” can be constructed. Each random list builds an edge of the hypernetwork. The edge in a hypernetwork is called a hyperedge since more than two vertices (words) can be contained in each edge. One important feature of the hypernetwork model is that a large number of random hyperedges are used to faithfully reproduce the observed data. Mathematically, the hypernetwork represents the joint probability of image *I* and text *T*:

$$P(I,T \mid W) = P(x_I, x_T \mid W) = P(x \mid W)$$

$$P(x \mid W) = \frac{1}{Z(W)} \exp\left[\sum_{k=1}^{K} \frac{1}{C(k)} \sum_{i_1,i_2,\ldots,i_k} w^{(k)}_{i_1 i_2 \ldots i_k} x_{i_1} x_{i_2} \ldots x_{i_k}\right]$$

where $x_{i_1} x_{i_2} \ldots x_{i_k}$ are hyperedges of order k. W are the parameters for the hypenetwork and $x = (x_I, x_T)$ is the training pattern consisting of image x_I and text x_T.

Once learned the joint probability of images and texts, the hypernetwork can be used to generate an image *I* given a text T by computing the conditional probability

$$P(I \mid T, W) = P(x_I \mid x_T, W) = \frac{P(x_I, x_T \mid W)}{P(x_T \mid W)}$$

where $P(x_T \mid W)$ can be estimated by counting the number of hyperedges matching with the text *T*. This probabilistic computation is inspired by the DNA computing-implemented hypernetwork model (Zhang & Kim, 2006) and performed by self-assembling the hyperedges associated with the given query.

Fig. 3 explains the retrieval process. Given a text query “leave my aura alone”, a large number of random hyperedges are sampled and matched with the hyperedges in the learned hypernetwork. The matching hyperedges have image parts associated with the text and the selected hyperedges are assembled to reconstruct an image. A similar scheme can be used to work in the reverse direction, i.e. generate a text from a given image.

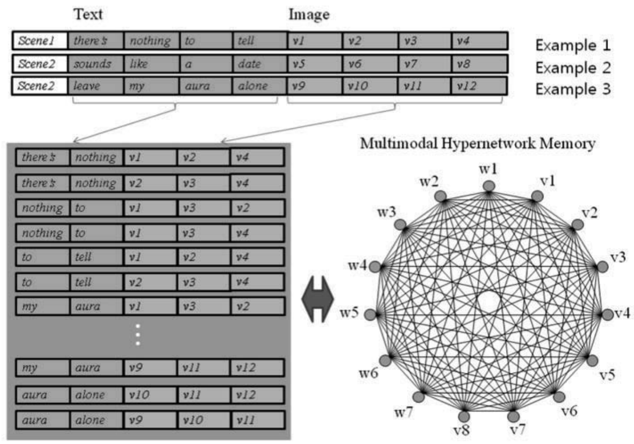

Fig. 2: Constructing a hypernetwork from a collection of text-image pairs. A large number of random fragments of the training data are sampled to make hyperedges constituting the hypernetwork. The hypernetwork is a parallel associative memory of a random hypergraph structure representing the joint probability of the text-image pairs.

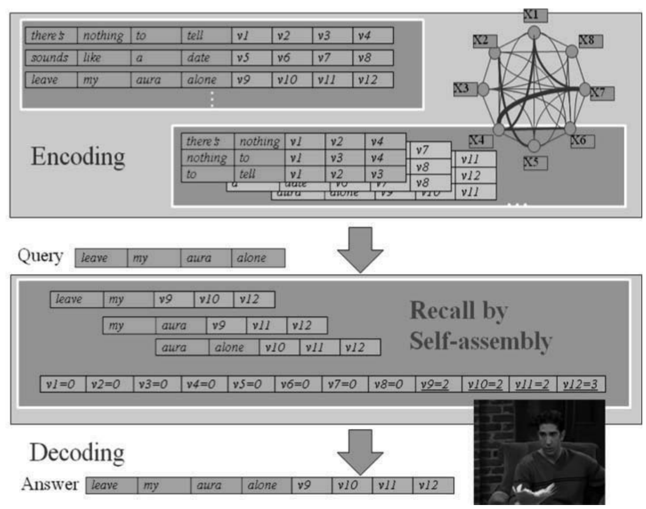

Fig. 3: Recall from the hypernetwork memory. From a given query (in this case, text), a large number of hyperedges are constructed and matched with the hyperedges stored in the hypernetwork. The matching hyperedges are then self-assembled to reconstruct the target image associated with the given query text.

Fig. 4 shows example results of text-to-image (T2I) experiments. The left column shows the query sentences and the middle column the images reconstructed by the hypernetwork. The right column shows the images retrieved using the reconstructed image as a visual query to the image database. It can be observed that the reconstructed images are blurred but very similar to the original video cuts. Using a similar procedure, the hypernetworks can produce texts from images.

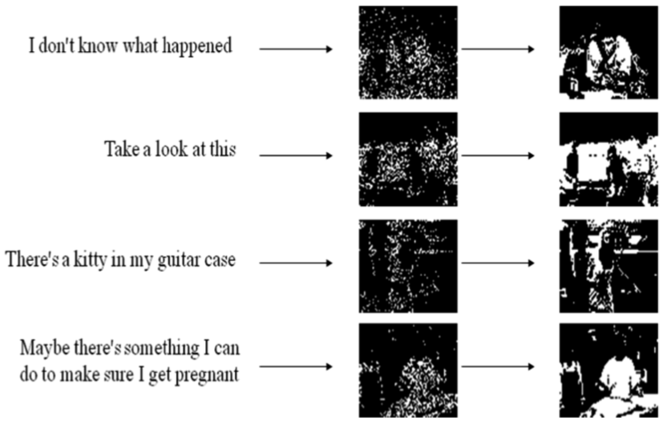

Fig. 4: Example results of text-to-image translation. Given a text query (left column), the hypernetwork assembles the hyperedges to reconstruct an image (middle column) associated with the text parts which matches with the query text. The generated "mental" images can be used to retrieve the original images (right column).

It should be noted that, though on a small scale, these experiments are non-trivial. In fact, no major machine learning algorithms, such as support vector machines or Bayesian networks, can solve these problems directly (Bishop 2006; Shawe-Taylor & Cristianini, 2004). The hypernetwork model is equipped with several properties that facilitate long-lasting learning in an online, dynamic environment.

The hypernetwork model has a *compositionality*, i.e. capable of incrementally building new concepts from primitive concepts. Smaller compositions of words represent global hyperfeatures while larger compositions represent more specialized, local hyperfeatures. Combining both types realizes the principle of *glocality*, i.e. a global representation made of a huge ensemble of localized micromodules, as found in population coding in brain circuits (Grillner & Graybiel, 2006). Though we did not go into much detail in this paper, the hypernetwork learning process starts with a large number of random compositions of the hyperedges and "evolves" the best-fitting compositions by randomly sampling and selecting them from the existing and new observations (text-image pairs) in sequence (Zhang, 2008b). This incremental restructuring or "self-organizing" process reflects the *continuity* principle, i.e. the new or frequently-observed data strengthen the hypernetwork memory, while the trace of the old data fades away as time goes on (unless they are observed again). More discussion on the importance of the shorter-term adaptability and the longer-term persistency in learning in a dynamic environment and, thus, for achieving human-level general intelligence can be found, for example, in (Langley, 2006; Zhang, 2008b).

Further Challenges for Cognitive Learning

The current implementation of the multimodal memory game can be extended to a more realistic lifelong learning situation where the players take turns in an arbitrary way and the learning process consists of a large number of sessions. This will offer several challenges for the machine learners. They can be divided into four categories.

1. Efficiency: How to process the learning material in real time
2. Capacity: How to manage the size of ever-growing memory storage
3. Accuracy: How to adapt to the environmental change for a short period
4. Consistency: How to maintain the consistency of learning for a long period

The traditional machine learning approach for handling these problems might be to design new learning algorithms. For example, one can attempt to devise algorithms to handle a bigger size of training sets more efficiently. Though necessary, this is not sufficient in everyday life situations since here the agent faces an unbounded size of data and the target system is dynamic. The agent should be more situation-aware and deal with the data explosion problem and the moving target problem.

Here we propose to use a cognitive approach, i.e. based on the findings and principles of learning and memory in human mind and brain (Roediger III *et al*., 2007; O'Reilly, 2006). Specifically, the learning strategy of the agent can be extended in the following dimensions.

1. Incremental learning: Learning is fundamentally incremental. New concepts are built upon existing concepts (Rumelhart & Norman, 1976).
2. Online learning: Once the learning material processed, it can be removed. Databasing all the training examples is practically impossible in many real-life situations.
3. Fast update: Learning should happen while performing the task. A fast, minimal update is sometimes more useful than a time-consuming fitting.
4. One-shot learning: Humans learn from a single example and so should the agent.
5. Predictive learning: The agent should be able to estimate the effect of learning.
6. Memory capacity: To deal with lifelong situations, we need a much bigger memory capacity than the usual memory size. Also important are forgetting and unlearning.
7. Selective attention: Humans have focus in its environment. Agents can also select data.
8. Active learning: The agent might ask questions to learn more effectively.
9. Context-awareness: Agents should be able to behave in a context-sensitive way.

10. Persistency: Learning does not mean just following a change, but it also pursuit a consistent long-term goal.
11. Concept drift: Learning in a dynamic environment should be able to track a moving target.
12. Multisensory integration: Multisensor data contains more information than single sensors. The agents should be able to integrate multiple sources of information (Zimmer *et al*., 2006).

Taking these into account, the MMG agent can learn, for example, by actively asking questions to the human players, selecting a subset of observed examples, and unlearning acquired knowledge to maintain the memory capacity within the limit.

Challenges for Human Teachers

Another approach to overcoming the difficulties of agents learning in everyday life situations comes from the human side. As a teacher to guide effective learning of the agents, humans have to deal with the following issues.

1. Efficiency: Reducing the data size.
2. Effectiveness: Providing more informative data.
3. Parameter setting: Making learning more efficient
4. Overhead: Decreasing teaching loads

Several specific ways are possible to achieve these general objectives. Here we discuss the challenges and strategies for the human teachers.

1. Getting the right feedback: To offer effective teaching signals, it is important to get the right feedback from the agent.
2. Sequencing the example: Different ordering of the same set of training examples may lead to improved results.
3. Identifying the weak points of the agent: Information gain is maximized by improving the weakest points.
4. Choosing the right problems: Teachers can choose different questions depending on the performance level of the agent.
5. Controlling the learning parameters: Setting the right parameters of the learning algorithm improves efficiency.
6. Evaluating the progress of learning: Teachers monitor the learner to measure its progress.
7. Estimating the difficulty of the problem: Knowing the level of the problem makes teaching more effective.
8. Generating new queries: Teachers design new problems to provide with the learner.
9. Modeling the effect of learning parameters: Teachers learn from interaction with the learners.
10. Catching the effect of environmental change: Knowledge about the environment leads to more informative teaching.
11. Maintaining minimal interactions: Human interactions involve time and effort on both learners and teachers. Only essential interactions should be made.
12. Multimodal interaction: Multiple modalities offer more effective communication between the teacher and the agent.

Learning from human teachers proved fruitful in many applications, including robotic agent domains (Thomaz & Breazeal, 2008; Chernova & Veloso, 2008). In cognitive science and psychometrics, techniques have been developed to generate test items (Embretson, 2002; Sternberg & Pretz, 2005). These are originally designed for human IQ tests, but can be equally used for evaluating and teaching agents. Teachers can also use query design algorithms. For example, genetic algorithms have been proposed to generate new examples and queries for teaching neural network learners (Zhang, 1993). Computational learning theory provides a mathematical basis for these query-based active learning algorithms (Valiant, 1984; Angluin 1988; Cohn *et al*., 1994).

Generally, teachers need to build a model of the learner and the environment. This model combined with meta-level learning strategies (Thrun & Pratt, 1998) can be used to provide more effective teaching signals to the agent. In case that many teachers offer teaching signals, as in the MMG game environment, the teachers can make a collective effort to offer teaching opportunities for the learning agents. This is analogous to the idea of human-based computation, which is defined as a computational process that performs its function via outsourcing certain steps to humans (Kosorukoff, 2001; von Ahn, 2006). This approach leverages differences in capabilities and alternative costs between humans and agents to achieve symbiotic human-agent interaction.

Concluding Remarks

The multimodal memory game (MMG) is used as an example of highly-interactive, lifelong learning scenario to address the limitations of current machine learning techniques. We discussed the challenges for future machine learning algorithms as well as the challenges for human teachers to deal with the problems of agents in everyday life situations. It is argued that the agents need to have cognitive learning capabilities, meaning that they require more attentive, active behavior rather than just parameter-fitting, passive adaptation. Also the human partners should play a more active role in teaching the agents by building models of the agent and the environment and by making necessary interactions with the agent. A merger of the *active learning* agents with *cognitively-aware* human teachers can result in *cognitive learning agents* that teach themselves. These *self-teaching agents* can attempt to design new queries and test their answers with initiated interaction with the human teachers.

Acknowledgments. The work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Pioneer Project (Phase I, 2008-2009) and the Korea Research Foundation (KRF) through the Fundamental Research Program (2008-2010).

References

Angluin, D. 1988. Queries and Concept Learning. *Machine Learning* 2: 319-342.

Bishop, C. M. 2006. *Pattern Recognition and Machine Learning*. New York, NY: Springer.

Chen, D. L., and Mooney, R. J. 2008. Learning to Sportscast: A Test of Grounded Language Acquisition. In *Proc. Int. Conf. on Machine Learning* (ICML-2008), 128-135.

Chernova, S., and Veloso, M. M. 2008. Teaching Multi-robot Coordination Using Demonstration of Communication and State Sharing. In *Proceedings of Autonomous Agents and Multiagent Systems* (AAMAS), 1183-1186.

Cohn, D., Atlas, L., and Ladner, R. 1994. Improving Generalization with Active Learning. *Machine Learning* 5(2): 201-221.

Embretson, S. E. 2002, Generating Abstract Reasoning Items with Cognitive Theory. In S. Irvine & P. Kyllonen (Eds.), *Item Generation for Test Development*, pp. 219-250, Mahwah, NJ: Erlbaum.

Gibbs, Jr. R.W. 2005. *Embodiment and Cognitive Science*. Cambridge, UK: Cambridge University Press.

Grillner, S. and Graybiel, A. M. eds. 2006. *Microcircuits: The Interface between Neurons and Global Brain Function*. Cambridge, MA: MIT Press.

Kosorukoff, A. 2001. Human-based Genetic Algorithm. *IEEE Transactions on Systems, Man, and Cybernetics* 5: 3464-3469.

Langley, P. 2006. Cognitive Architectures and General Intelligent Systems. *AI Magazine*, 33-44.

MacLeod, C. M., Hunt, E. B., & Mathews, N. N. 1978. Individual Differences in the Verification of Sentence-Picture Relationships, *Journal of Verbal Learning and Verbal Behavior* 17: 493-507.

O'Reilly, R. C. 2006. Biologically Based Computational Models of High-level Cognition. *Science* 314: 91-94.

Roediger III, H. L., Dudai, Y., and Fitzpatrick, S. M. 2007. *The Science of Memory: Concepts*. Oxford, UK: Oxford University Press.

Rumelhart, D. E. and Norman, D. A. 1976. Accretion, Tuning and Restructuring: Three Modes of Learning. In J. W. Cotton and R. Klatzky (Eds.), *Semantic Factors in Cognition*. Lawrence Erlbaum.

Shawe-Taylor, J. and Cristianini, N. 2004. *Kernel Methods for Pattern Analysis*. Cambridge, UK: Cambridge University Press.

Sternberg, R. J., and Pretz, J. E., eds. 2005. *Cognition and Intelligence: Identifying the Mechanisms of the Mind.* Cambridge, UK: Cambridge University Press.

Thomaz, A. L., and Breazeal, C. 2008. Teachable Robots: Understanding Human Teaching Behavior to Build More Effective Robot Learners. *Artificial Intelligence* 172(6-7): 716-737.

Thrun, T. and Pratt, L.Y. eds. 1998. *Learning To Learn*. Amsterdam, Holland: Kluwer.

Valiant, L. G. 1984. A Theory of the Learnable. *Communications of the ACM*, 27(11):1134-1142.

von Ahn, L. 2006. Games with a Purpose. *IEEE Computer*, 96-98.

Zhang, B.-T. 2008a. Cognitive Learning and the Multimodal Memory Game: Toward Human-Level Machine Learning. In *IEEE World Congress on Computational Intelligence (WCCI-2008),* Special Session on Cognitive Architectures: Towards Human-Level Intelligence. Piscataway, NJ: IEEE Press.

Zhang, B.-T. 2008b. Hypernetworks: A Molecular Evolutionary Architecture for Cognitive Learning and Memory, *IEEE Computational Intelligence Magazine* 3(3): 49-63.

Zhang, B.-T. 1993. Teaching Neural Networks by Genetic Exploration. *Arbeitspapiere der GMD*, No. 805, German National Research Center for Computer Science (GMD), St. Augustin/Bonn.

Zhang, B.-T. and Kim, J.-K. 2006. DNA Hypernetworks for Information Storage and Retrieval. *Proc. 2006 Int. Annual Meeting on DNA Computing* (*DNA12*), *LNCS* 4287: 298-307.

Zimmer, H. Z., Mecklinger, A., and Lindenberger, U. 2006. Levels of Binding: Types, Mechanisms, and Functions of Binding in Memory. In Zimmer, H. Z., Mecklinger, A., and Lindenberger, U. (Eds.), *The Handbook of Binding and Memory.* Oxford, UK: Oxford University Press.