Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Personal View:

Learning deep IA bidirectional intelligence*

Lei XU[1,2]

[1]*Centre for Cognitive Machines and Computational Health (CMaCH), School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

[2]*Neural Computation Research Centre, Brain and Intelligence Sci-Tech Institute, Zhangjiang Lab, Shanghai 201210, China*

E-mail: lxu@cs.sjtu.edu.cn

Received Sept. 30, 2019; Revision accepted Dec. 15, 2019; Crosschecked Dec. 23, 2019

Abstract: There has been a framework sketched for learning deep bidirectional intelligence. The framework has an inbound that features two actions: one is the *acquiring* action, which gets inputs in appropriate patterns, and the other is A-S cognition, derived from the abbreviated form of words *abstraction and self-organization*, which abstracts input patterns into concepts that are labeled and understood by self-organizing parts involved in the concept into structural hierarchies. The top inner domain accommodates relations and a priori knowledge with the help of the *A-I thinking* action that is responsible for the accumulation-amalgamation and induction-inspiration. The framework also has an outbound that comes with two actions. One is called *I-S reasoning*, which makes inference and synthesis (I-S) and is responsible for performing various tasks including image thinking and problem solving, and the other is called the *interacting* action, which controls, communicates with, and inspects the environment. Based on this framework, we further discuss the possibilities of design intelligence through synthesis reasoning.

Key words: Abstraction; Least mean square error reconstruction (Lmser); Cognition; Image thinking; Abstract thinking; Synthesis reasoning

https://doi.org/10.1631/FITEE.1900541 **CLC number:** TP18

1 Synthesis reasoning

The first wave of artificial intelligence in China began in the early 1980s. During that time, studies focusing on symbolic reasoning dominated the entire international literature of artificial intelligence. In contrast, Qian (1983) believed that image thinking plays a leading role in intelligence and that study on image thinking should be a future breakthrough point, which has been followed and developed by Pan (1996). Pan believed that reasoning research started gradually from deductive logic to visual reasoning and demonstrated such a loosening tendency in the reasoning process, leading reasoning research to thinking simulation. Also, Pan proposed a synthesis reasoning model, which expounds the relationship between image thinking and traditional reasoning and compares the characteristics of image thinking with traditional reasoning.

Following Pan (1996), Pan's team has not only explored synthesis reasoning based on a single source or multiple sources, but also tried several applications in creativity and design to examine the possible point set in a synthesis space specified by one or several known cases that correspond to examining possible design sketches or prototypes accordingly. Undoubtedly, synthesis reasoning plays an important role in the study of design intelligence.

Unfortunately, like that building a traditional expert system bases on a man-crafted rule base, an

* Project supported by the National New Generation Artificial Intelligence Project, China (No. 2018AAA0100700) and the Zhiyuan Chair Professorship Start-up Grant from Shanghai Jiao Tong University, China (No. WF220103010)

ORCID: Lei XU, https://orcid.org/0000-0002-2752-1573

implementation of Pan (1996) relies on a synthesis space specified by man-crafted sources, parts, structures, and fields, which hinders the further development and applications. Nowadays, such a direction may be reactivated by recent advances on bidirectional deep learning.

Here we outline a proposal for learning deep yIng-yAng bidirectional intelligence (IA-BI) that performs cognition and problem solving with image thinking, abstract thinking, and creative thinking, with further details referred to a recent overview in Xu (2019b). The subsequent two sections briefly summarize and elaborate this deep IA-BI as illustrated in Fig. 1, echoing the views of Qian (1983) and developing synthesis reasoning such that the synthesis space is learned from data, and reasoning is driven by data or cause, or both.

2 A-S cognition and image thinking

We begin by observing Fig. 1c. After doing the action of *acquiring* information from data to form a vector or image $\boldsymbol{X}$, we perform the action of *abstraction and self-organization* that abstracts $\boldsymbol{X}$ into a much compressed code via a deep neural network by supervised learning based on a set of paired samples of $\boldsymbol{X}$.

In the inner coding domain (shortly, I-domain), the code is either a label v_i that denotes a pattern, a class, a concept, or a number of attributes $\boldsymbol{Y}_i = [y_i^{(1)}, y_i^{(2)}, \ldots, y_i^{(m)}]^{\mathrm{T}}$ that describe the pattern, class, and concept as a preliminary cognition. Also, as suggested in Xu (2019a), there may be an icon or primitive as the third coding type that encodes the structural and topological property of $\boldsymbol{X}$. However, aiming at merely an abstraction mapping, such a top-down supervised learning just performs a preliminary cognition.

Another preliminary function of cognition comes from a bottom-up unsupervised learning as illustrated in Fig. 1d, which cascades $\boldsymbol{X} \to \boldsymbol{Y}$ with an inverting or inferring mapping $\boldsymbol{Y} \to \boldsymbol{X}$ (I-mapping, for short) to generate one $\hat{\boldsymbol{X}}$ as a reconstruction of $\boldsymbol{X}$, such that $\boldsymbol{X} \to \boldsymbol{Y} \to \hat{\boldsymbol{X}}$ approximates an identical mapping. One earliest example is the autoencoder (AE) proposed by Ballard (1987) that cascades

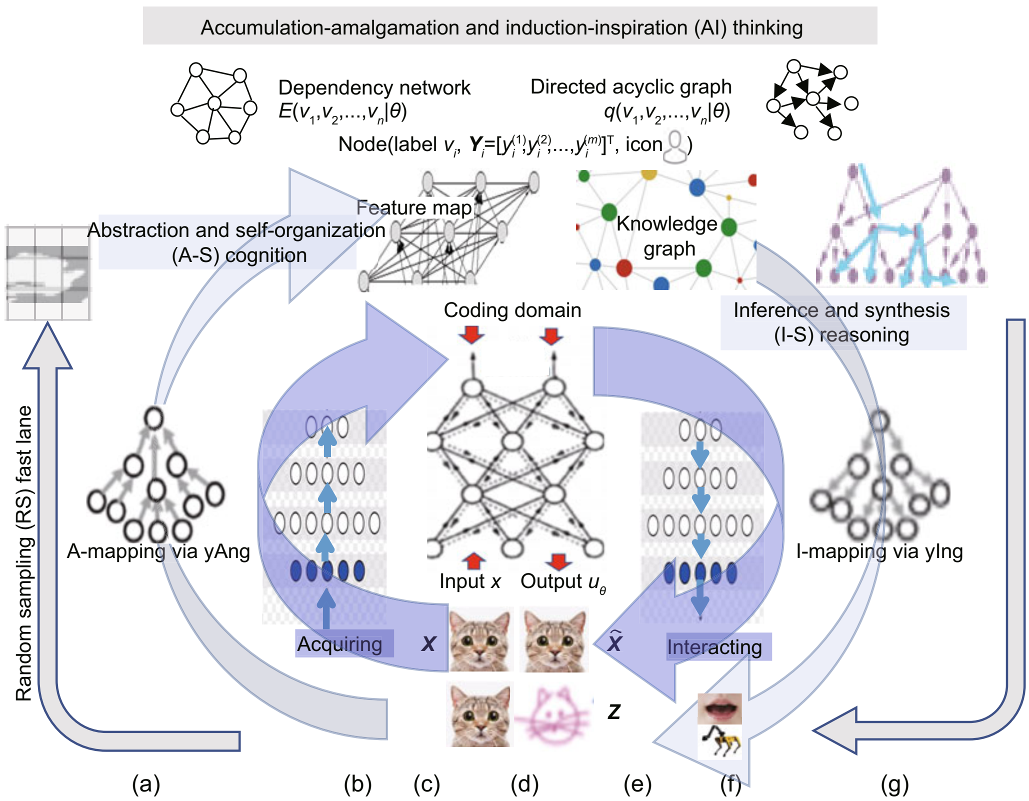

Fig. 1 Deep yIng-yAng bidirectional intelligence (IA-BI)

two networks in a same architecture as illustrated in Figs. 1c and 1e, in which all unknown parameters are determined by the least mean square error (MSE) $E\|\boldsymbol{X} - \hat{\boldsymbol{X}}\|^2$. Here, this unsupervised learning perceives $\boldsymbol{X}$ by a much compressed code $\boldsymbol{Y}$.

Initially proposed by Xu (1991, 1993), the least mean square error reconstruction (Lmser) self-organization overlaps the two networks in Figs. 1c and 1e into the one at the center of Fig. 1d, making the situation significantly different.

As illustrated in Fig. 2a, one layer Lmser performs self-organization (Xu, 1991, 1993) such that the downward reconstruction u_θ approximates the input $\boldsymbol{x}$, the weights of each unit learn a feature template that detects a particular orientation, and the upward mapping performs independent component analysis (ICA), i.e., making the upper layer units become mutually independent.

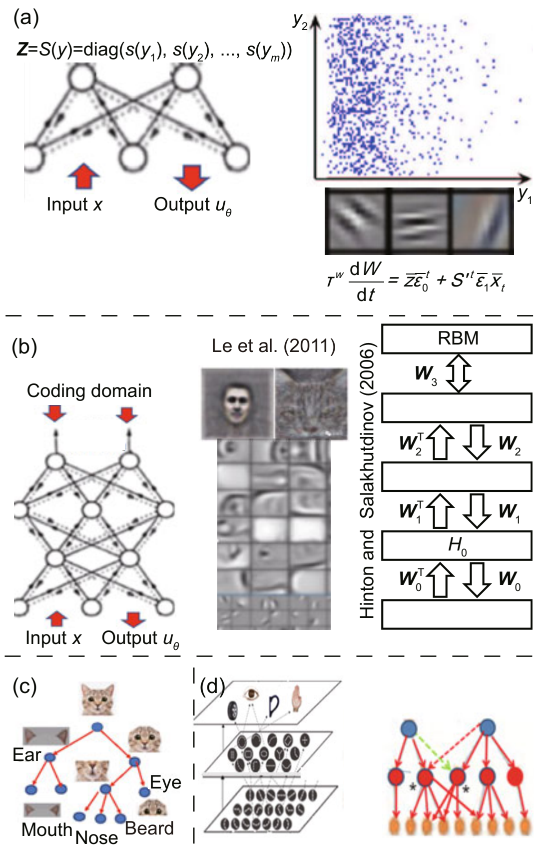

Fig. 2 Lmser facilitates to form concept hierarchies

Recalling Sections 5(b) and 5(c) of Xu (1991), it was also speculated that multilayer Lmser self-organization makes the higher layers develop into templates of higher-order features, while the top layer forms certain concepts, which was echoed and further developed in Hinton and Salakhutdinov (2006). As illustrated in Fig. 2b, they experimentally showed multilayer developments of feature templates by stacked restricted Boltzmann machines (RBMs) that are quite similar to those of Lmser, with details referred to Fig. 4 in Xu (2019b). Subsequently, Le et al. (2011) found that the head and cat face emerge via their computation in Google.

Following the well-known feature detection theory by Hubel and Wiesel (1962), it is understood that a concept is an organized hierarchy of parts or components that form the concepts, as illustrated in Fig. 2d. The promising nature of such a hierarchy is that the son nodes become independent once their common father node is given. Thus, as the supervision propagates from the top down to the bottom, parts or components become mutually independent layer by layer. Unfortunately, supervised learning by back-propagation cannot enable this nature.

Promisingly, it follows from ICA, which is illustrated in Fig. 2a, that a multilayer Lmser self-organization attempts to make units per layer become mutually independent from the bottom up layer by layer, thus making concepts or components become easier to organize. Together, the bottom-up Lmser self-organization facilitates in forming such hierarchies, while the top-down supervised learning helps reallocate the hierarchies such that common subtrees become effectively shared, as illustrated in Fig. 2d.

Moreover, the overlap of two networks in Figs. 1c and 1e into the one at the center of Fig. 1d also results in several duality natures and favorable characteristics, featuring not only skip connections like those recently popularized U-Net, ResNet, and DenseNet, but also feedback connections like those in recurrent networks. Refer to Section 2 of Xu (2019b) for further details.

Furthermore, as suggested in Xu (2019a), a random sampling (RS) fast lane, as illustrated in Fig. 1a, obtains information by RS from the input pattern and quickly delivers it to a higher or top layer to form samples of low-resolution images, in which some large-scale feature or topological nature can be cognized and then propagated downward as attention to coordinate the bottom-up self-organization.

Targeting at minimizing $E\|\boldsymbol{X} - \hat{\boldsymbol{X}}\|^2$, both AE and Lmser have not considered a priori knowledge about $\boldsymbol{Y}$ that is typically described by a distribution $q(\boldsymbol{Y})$ in the I-domain. Denoting the I-mapping generally by $q(\boldsymbol{X}|\boldsymbol{Y})$, the task of reconstructing $\boldsymbol{X}$ is made by the maximum likelihood (ML) on $q(\boldsymbol{X}) = \int q(\boldsymbol{X}|\boldsymbol{Y})q(\boldsymbol{Y})\mathrm{d}\boldsymbol{Y}$, toward which the A-mapping needs to be its posterior $p(\boldsymbol{Y}|\boldsymbol{X}) = q(\boldsymbol{X}|\boldsymbol{Y})q(\boldsymbol{Y})/q(\boldsymbol{X})$ that is often computationally intractable. Proposed by Dayan et al. (1995), the Helmholtz machine and variational learning approximate ML by a $p(\boldsymbol{Y}|\boldsymbol{X})$ in a prespecified structure.

About the same period, Xu (1995) proposed Bayesian Ying-Yang (BYY) learning that considers either best matching or harmony between two different models of the joint distribution of $\boldsymbol{X}$ and $\boldsymbol{Y}$. One is $q(\boldsymbol{X}|\boldsymbol{Y})q(\boldsymbol{Y})$, and the other is $p(\boldsymbol{Y}|\boldsymbol{X})p(\boldsymbol{X})$ with $p(\boldsymbol{X})$ for the samples of $\boldsymbol{X}$ from the actual world or shortly A-domain. Following the ancient yIng-yAng (IA) philosophy ("Ying" is spelled "Yin" in the current Chinese Pin Yin system that could be backtracked over 400 years from the initiatives of M. Ricci and N. Trigault. However, the length of "Yin" loses its harmony with Yang; thus, "Ying" is preferred since 1995 (Xu, 1995).), a visible domain is called the yAng domain, while the invisible domain is called the yIng domain, which both coincide with the A-domain and I-domain, respectively. Also, an I-mapping from an inner DNA to a real body coincides with the role of a yIng animal, and an A-mapping from a real body to an inner DNA coincides with the role of a yAng animal. With $q(\boldsymbol{X}|\boldsymbol{Y})$ for the I-mapping and $p(\boldsymbol{Y}|\boldsymbol{X})$ for the A-mapping, BYY learning also shares with the basic spirit of the ancient Chinese IA concept. Readers are referred to Xu (1995, 2010) for details about BYY learning and to Section 3 of Xu (2019b) for its relations to ML and variational learning.

Generally, a joint implementation of I-mapping and A-mapping may be extended to $\boldsymbol{X} \to \boldsymbol{Y} \to [\hat{\boldsymbol{X}}, \boldsymbol{Z}]$ that performs various transformations from one pattern $\boldsymbol{X}$ to the other $\boldsymbol{Z}$, as illustrated at the bottom row in Fig. 1d, such as image to image, language to language, text to image, text to sketch, sketch to image, 2D image to 3D image, image to sentence, music to dance, and past to future, all of which are examples of image thinking, echo the views of Qian (1983), and share with the main features of synthesis reasoning by Pan (1996).

3 I-S reasoning vs. synthesis reasoning

Image thinking is characterized by simultaneously implementing A-mapping and I-mapping, while I-mapping is identified by performing either a statistical inference $\boldsymbol{Y} \to \boldsymbol{Z}$ holistically via a neural network or a synthesis reasoning that synthesizes $\boldsymbol{Z}$ from $\boldsymbol{Y}$ according to the hierarchical structures learned for A-S cognition, but inversely along the outward direction. Accordingly, we abbreviate the term of such an inference and synthesis (I-S) process into *I-S reasoning*. Also, instead of outputting the pattern $\boldsymbol{Z}$, the process may perform sequential decision with $\boldsymbol{Z}$ consisting of a sequence of a label z_t that indicates each choice. Alternatively, the process may produce a reasoning tree that proves a statement.

This *I-S reasoning* action is followed by the *interacting* action, which not only controls and communicates with the counterparts in the external world, but also examines whether outcomes are good enough. Also, communication may directly come from A-I thinking through a shortcut, as illustrated in Fig. 1g.

This I-S reasoning action is driven either directly by the outcomes of A-S cognition for performing the mapping $\boldsymbol{X} \to \boldsymbol{Y} \to [\hat{\boldsymbol{X}}, \boldsymbol{Z}]$ or indirectly via the inner I-domain that accommodates relations and a priori knowledge, featured by an integrated action named A-I thinking that makes accumulation-amalgamation and induction-inspiration (A-I), as illustrated on the top of Fig. 1. Specifically, evidence is accumulated to enhance or weaken the outcomes of cognition. Also, cognitions and common knowledge (e.g., feature map and knowledge graph) are amalgamated or united to form new concepts and analyze dependence among nodes. Moreover, accumulation-amalgamation may incur induction that discovers causal relations to obtain a directed acyclic graph (DAG). Occasionally, there are certain inspirations that may emerge to drive *I-S reasoning*.

This *I-S reasoning* action provides opportunities and possibilities of reactivating the study on synthesis reasoning. In sequel, we give some discussions related to both large- and small-sized samples.

In a circumstance of a large number of samples about cases involved in a design task, we obtain a learned deep IA-BI system as illustrated in Fig. 1, in which the I-domain acts as the synthesis space considered in Pan (1996), while the synthesis function $\mathrm{SS}(x, y, z)$ given by Definition 3 in Pan (1996) is generalized here by the I-mapping $\boldsymbol{Y} \to \boldsymbol{Z}$ that actually performs a synthesis function $Z = \mathrm{SS}(\boldsymbol{Y})$.

Synthesis reasoning is considered along two possible directions. First, interpolation is made by considering $\mathrm{SS}(Y_s)$ with Y_s coming from a linear or nonlinear combination of a number of codes (e.g., Y_A, Y_B, Y_C) obtained directly from the learned A-mapping $\boldsymbol{X} \to \boldsymbol{Y}$. Second, synthesis reasoning $\mathrm{SS}(Y_s)$ considers Y_s resulted from A-I thinking on the knowledge graph, dependency network, and DAG.

In a circumstance of few exemplars, we may consider transfer learning. First, we learn a deep IA-BI system from a related task, with Lmser self-organization and supervised learning jointly obtaining the hierarchies as illustrated in Fig. 2d for I-mapping. Second, we prune the extra links in the hierarchies using causal analysis and knowledge from few exemplars and then make an interpolation under the constraints of hierarchies.

Based on knowledge from the exemplars, we may even consider an architecture illustrated in Fig. 2c as the architecture of Lmser, with each node outputting multiple attributes not only to indicate the node's activating level but also to describe the major features of the corresponding component. That is, each node acts as a vector function of multiple vectors. The function form may be a man-crafted one according to the exemplar structure (e.g., the structure of cat face), and unknown parameters are learned jointly by Lmser self-organization and supervised learning.

Compliance with ethics guidelines

Lei XU declares that he has no conflict of interest.

References

Ballard DH, 1987. Modular learning in neural networks. Proc 6th National Conf on Artificial Intelligence, p.279-284.

Dayan P, Hinton GE, Neal RM, et al., 1995. The Helmholtz machine. *Neur Comput*, 7(5):889-904.

Hinton GE, Salakhutdinov RR, 2006. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504-507. https://doi.org/10.1126/science.1127647

Hubel DH, Wiesel TN, 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*, 160(1):106-154.

Le QV, Ranzato M, Monga R, et al., 2011. Building high-level features using large scale unsupervised learning. https://arxiv.org/abs/1112.6209

Pan Y, 1996. The synthesis reasonning. *Patt Recogn Artif Intell*, 9:201-208.

Qian X, 1983. On thinking sciences. *Chin J Nat*, 8:566.

Xu L, 1991. Least MSE reconstruction for self-organization: (II) further theoretical and experimental studies on one-layer nets. Proc IEEE Int Joint Conf on Neural Networks, p.2362-2367.

Xu L, 1993. Least mean square error reconstruction principle for self-organizing neural-nets. *Neur Netw*, 6(5):627-648.

Xu L, 1995. YING-YANG machines: a Bayesian-Kullback scheme for unified learning and new results on vector quantization. Proc Int Conf on Neural Information Processing, p.977-988.

Xu L, 2010. Bayesian Ying-Yang system, best harmony learning, and five action circling. *Front Electr Electron Eng China*, 5(3):281-328.

Xu L, 2019a. Deep IA-BI and five actions in circling. Int Conf on Intelligent Science and Big Data Engineering, p.1-21.

Xu L, 2019b. An overview and perspectives on bidirectional intelligence: Lmser duality, double IA harmony, and causal computation. *IEEE/CAA J Autom Sin*, 6(4):865-893. https://doi.org/10.1109/JAS.2019.1911603