A Survey on Autonomous Learning

Jiaqi Xi

July 2019

Introduction

- We develop this survey in four parts.

    1. What is autonomous learning? What are the characteristics of autonomous learning?
    2. What are the related keywords of autonomous learning?
    3. How should autonomous learning be implemented?
    4. What are the difficulties and challenges realizing autonomous learning?

Part 1

What is autonomous learning?

What are the characteristics of autonomous learning?

Part 1

- Autonomous learning

— how an embodied agent can determine what to learn, how to learn, and how to judge the learning success.

(Note: This descriptive definition is extracted from https://al.is.tuebingen.mpg.de/, a group focusing on autonomous learning. A more specific definition involves the characteristics of autonomous learning.)

Part 1

- [1] Characteristics — what to learn
  - **Attention**:

    ① **Familiar** subjects: Recall long-term memory and select the most important part to focus on.

    ② **Unfamiliar/new** subjects: Spark curiosity. Able to monitor its learning progress and orient resources toward problems that maximize information gain.

Part 1

- [1] Characteristics — how to learn
  - **Complex decision making**: Compare different paths and select the most suited one. Summarize and memorize some already made decisions for further processing.
  - **Global availability**:

    ① **Internal**: The knowledge acquired and integrated can be held over time; thus able to globally recall or reuse. (similar to experience and memory)

    ② **External**: Aware of others' information. Able to share information and collaborate.

Part 1

- [1] Characteristics — how to judge the learning success
  - **Introspection (self-aware)**: Able to represent the extent and limits of the machine itself, namely develop an understanding of itself. Know the facts that others may have a different viewpoint, thus ready to share information and re-describe itself. With the help of curiosity, able to self-organize the learning curriculum, self-generate and self-select the goals.

Part 2

What are the related keywords of autonomous learning?

Part 2

- [3] **Lifelong learning**:
    - Compared with the frozen and deployed trained models in traditional machine learning, lifelong learning studies continual learning across tasks and data.
    - The goal is to overcome catastrophic forgetting and **accumulate knowledge across tasks** (typically via model sharing), thus realizing **further adaptation or customization** of one single model.

Part 2

- **Meta learning**:
  - 'Meta' stands for the kernel knowledge of a set of learning tasks. The goal is to design models that can learn new skills or rapidly adapt to new environments with minimal training examples by efficiently extracting the 'meta' knowledge.
- [5] **Reinforcement learning**:
  - Studies how an agent can maximize its cumulative **reward** in a previously **unknown environment**, which it learns about through experience. It is closely related to 'curiosity'.

Part 2

- **Small Sample Learning**:
    - Learn new concepts or tackle novel tasks from small sample with strong generalization ability.
- **Multi-agent Reinforcement Learning**:
    - The collaboration and competition between both homogeneous and heterogeneous agents.
- **Self-aware**

Part 3

How should autonomous learning be implemented?

Part 3

Attention—Unfamiliar/new Subjects

[5] **VIME**(Variational Information Maximizing Exploration): A curiosity-driven exploration strategy for continuous control tasks.

① Using information gain as intrinsic rewards.

② Variational inference is used to approximate the posterior distribution of a Bayesian neural network that represents the environment dynamics.

Part 3

Attention—Familiar Subjects & Global Availability—Internal

[3] **MAS**(Memory Aware Synapses): Based on a trained model and selectively preserve the parameters considering their importance.

① 'how to learn'(internal global availability): Not only memorize the trained model and the parameters, but also stratify the parameters and estimate the **importance weights** by observing the **sensitivity** of outputs.

② 'introspection': Allow for **adaption** to specific conditions and **continuous updating** of importance weights.

③ 'what to learn': **Penalize** the changes of important parameters by regularizing the loss function of new tasks.

Part 3

Attention—Familiar Subjects & Global Availability—Internal

[3] Add a perturbation to test the sensitivity:

$$F(x_k;\theta+\delta) - F(x_k;\theta) \approx \sum_{i,j} g_{ij}(x_k)\delta_{ij}$$

Here, function g is the **gradient** of F with respect to parameter θ. When F is multi-dimensional, use the gradients of the squared **L2 norm** of the output:

$$g_{ij}(x_k) = \frac{\partial[\ell_2^2(F(x_k;\theta))]}{\partial\theta_{ij}}$$

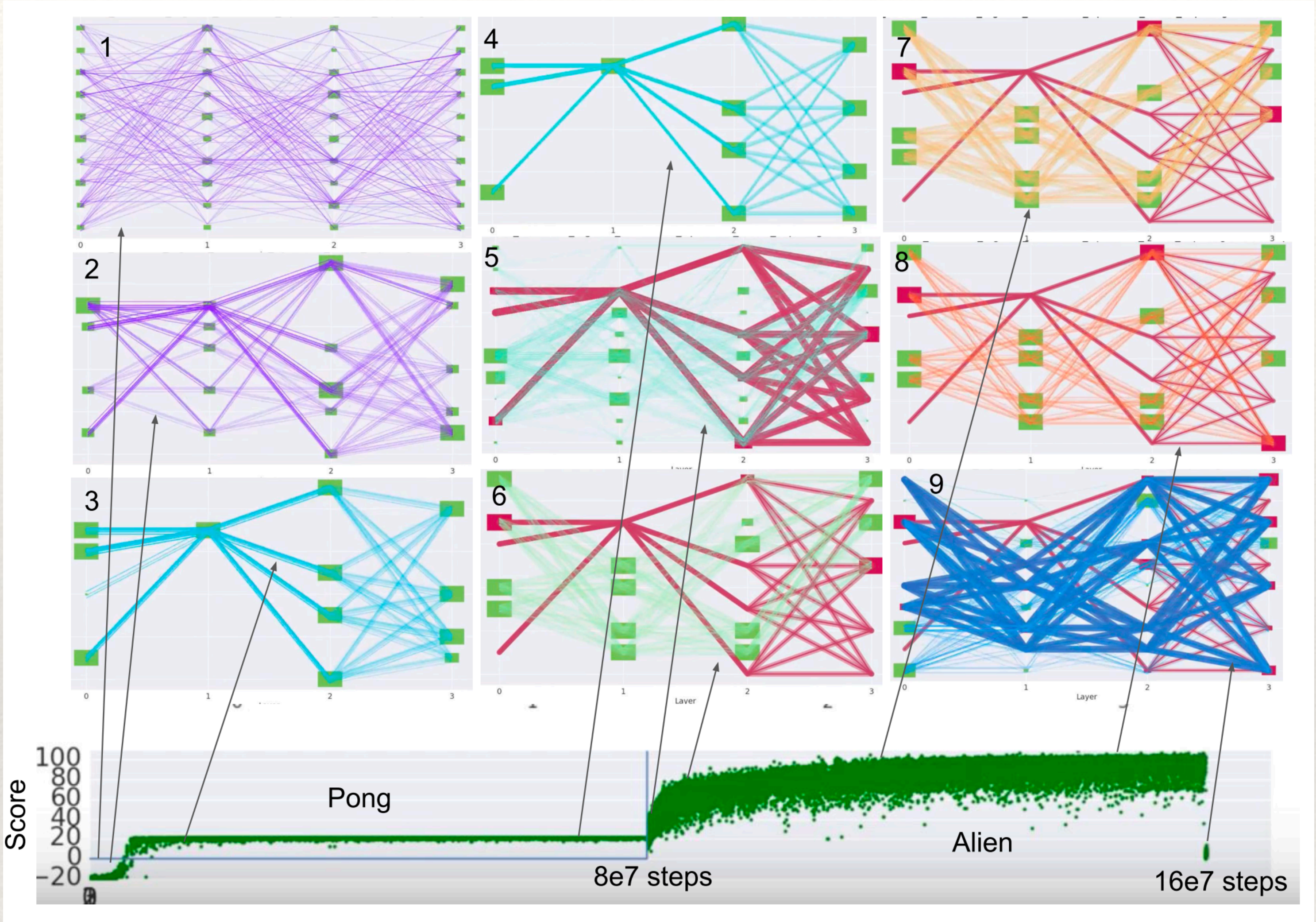
Then we have the importance weight Ω and loss function (here θ* are old parameters):

$$\Omega_{ij} = \frac{1}{N}\sum_{k=1}^{N} \| \, g_{ij}(x_k) \, \|$$

$$L(\theta) = L_n(\theta) + \lambda \sum_{i,j} \Omega_{ij}(\theta_{ij} - \theta_{ij}^*)^2$$

Part 3

Attention—Familiar Subjects & Global Availability—Internal



[6] **Figure: Pathnet**: The whole graph is fixed. When tasks change, the model only needs to find some new, suitable paths for the current task.

Part 3

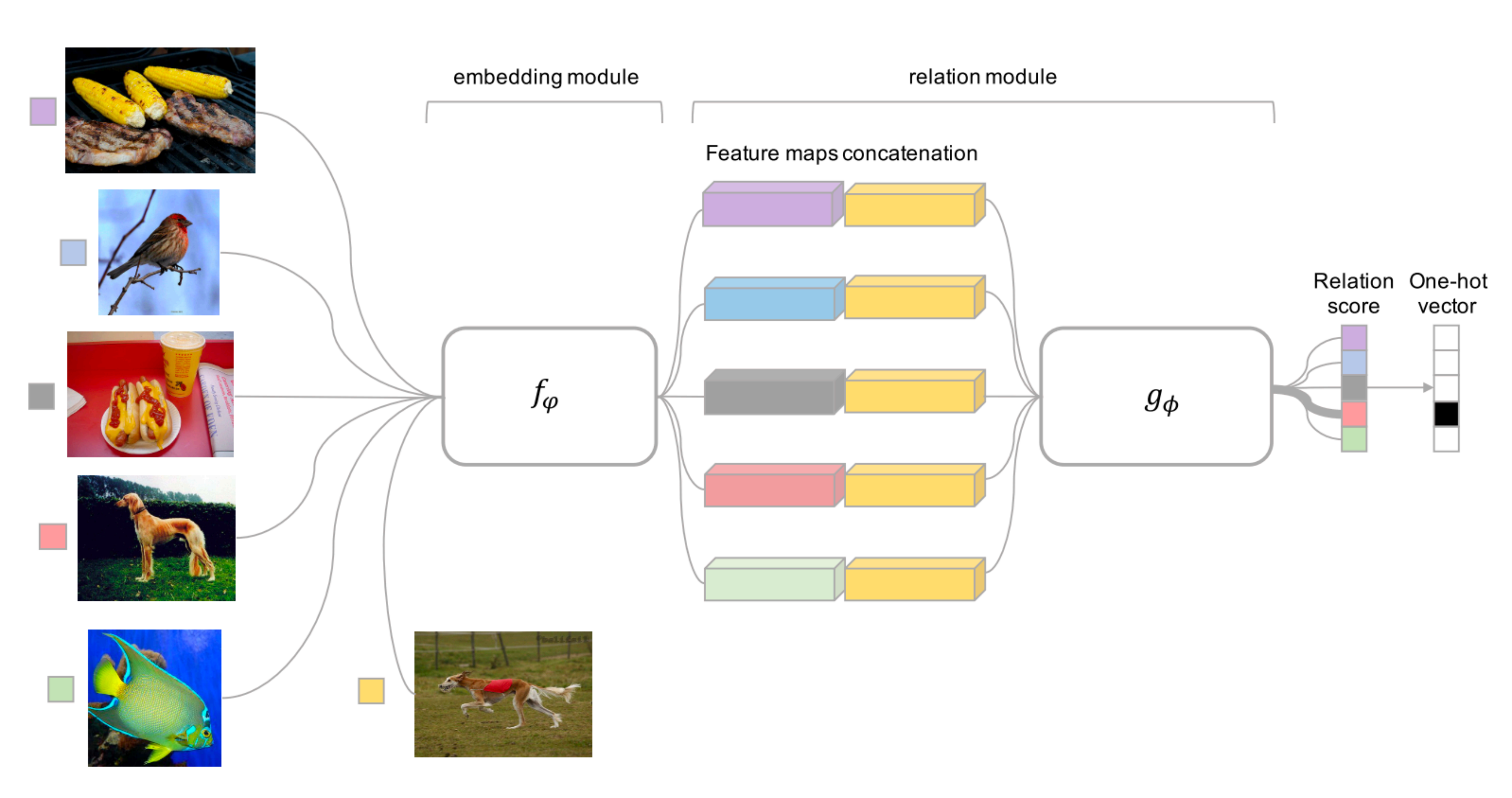
How to learn—Meta Learning

[7] **Relation Network** for Few-shot Learning:

- a **two-branch** Relation Network(RN) learning to compare query images against few-shot labeled sample images
- Of the two branches, one is the *embedding module* which generates **representations** of the query and training images, the other is *a relation module* that helps find the matching category by **similarity**.

Part 3

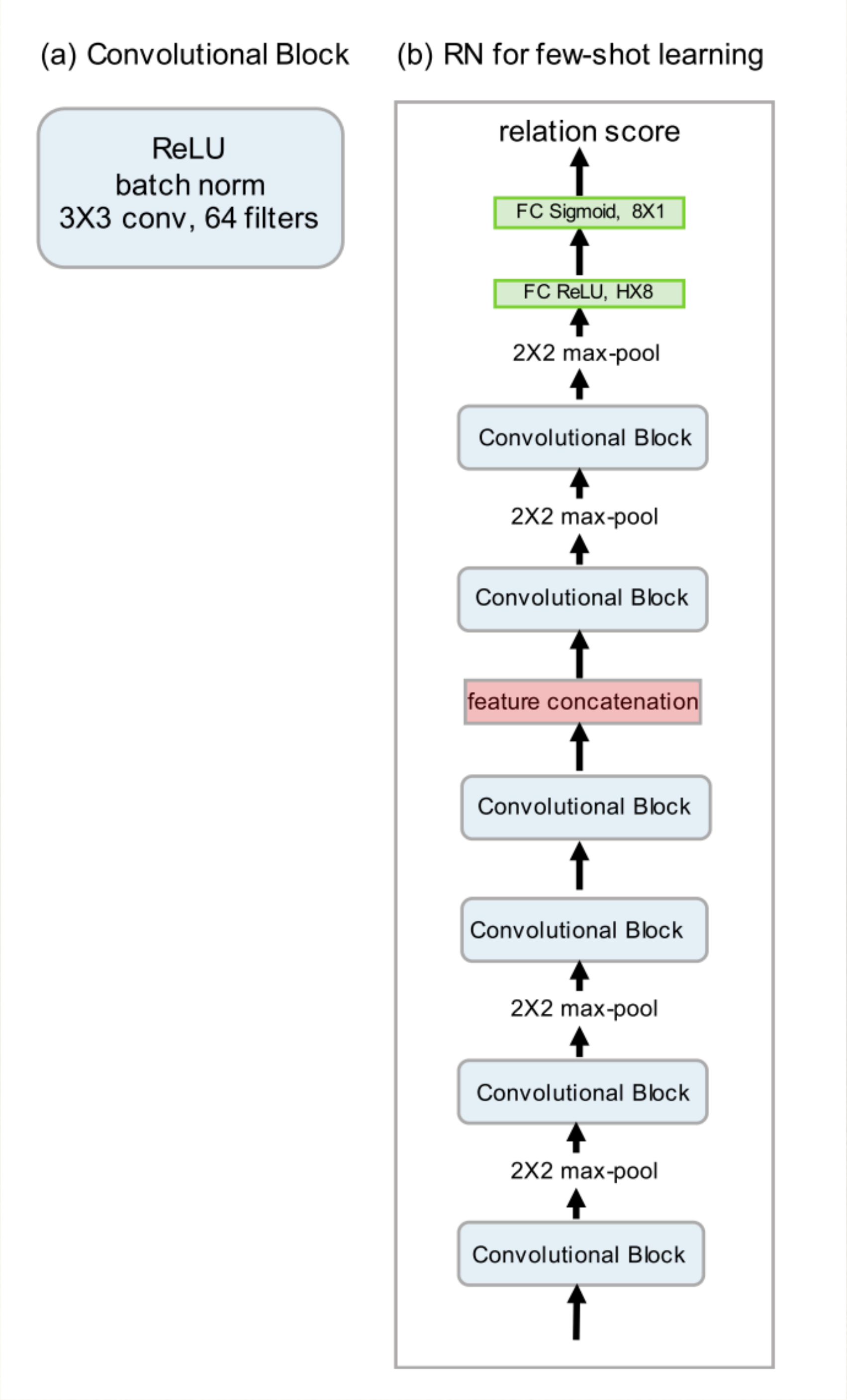
How to learn—Meta Learning



[7] **Figure:** Relation Network architecture for a 5-way 1-shot problem with one query example.

Part 3

How to learn—Meta Learning



- **Advantage**:

Compared with fixed features and fixed metrics, this method learn features and similarity metrics simultaneously, relying less on external help.

[7] **Figure:** Relation Network architecture for few-shot learning.

Part 3

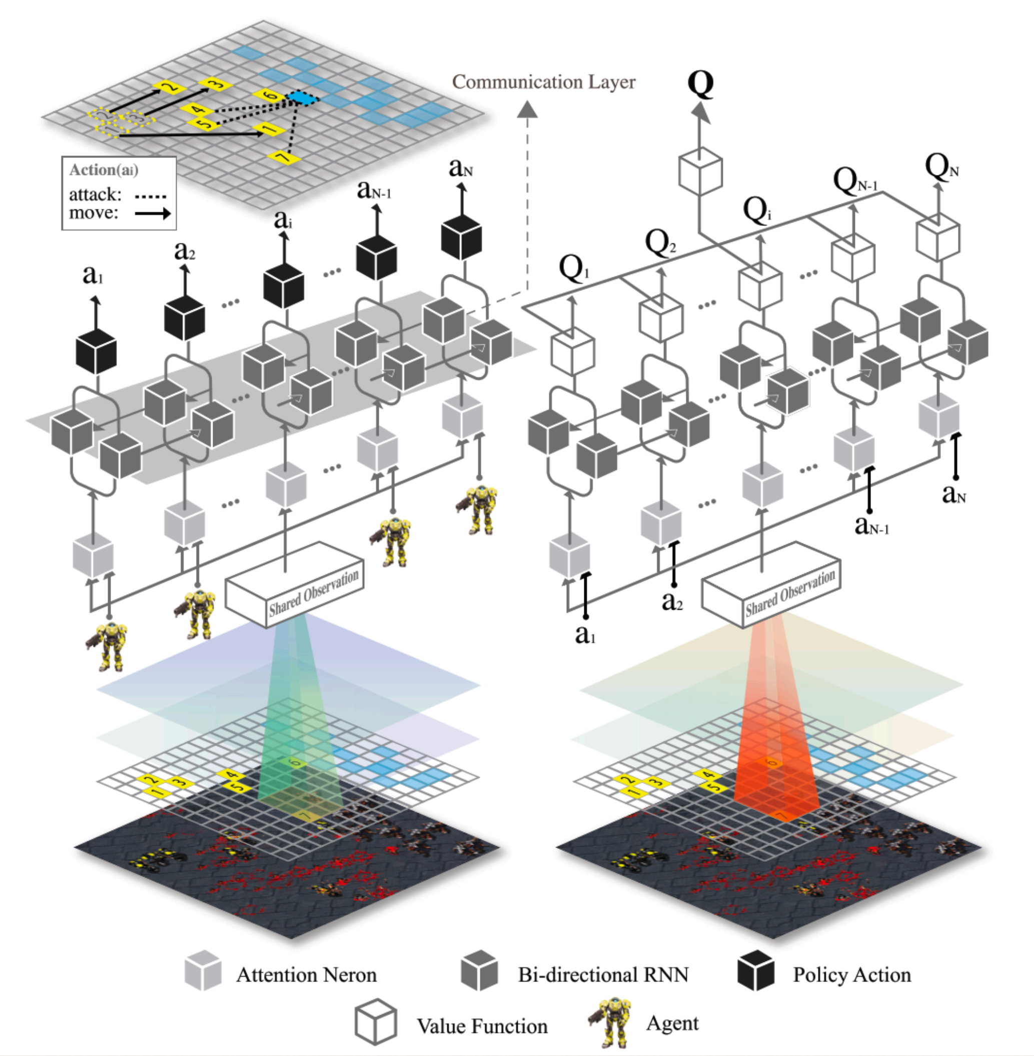
Global Availability—External

[9][10] **Multiagent Bidirectionally Coordinated Network**:

- a method of multi-agent communication
- uses parameter sharing
- independent rewards
- Constructs a vectorized actor-critic framework, where each dimension corresponds to an agent. The coordination is done by bi- directional recurrent communications in the **internal(hidden)** layers.

Part 3

Global Availability—External



[10] **Figure:** (i) communicate with bidirectional RNN; (ii) learning is done by multi-agent deterministic actor-critic; (iii) share parameters between homogeneous agents to ensure scalability

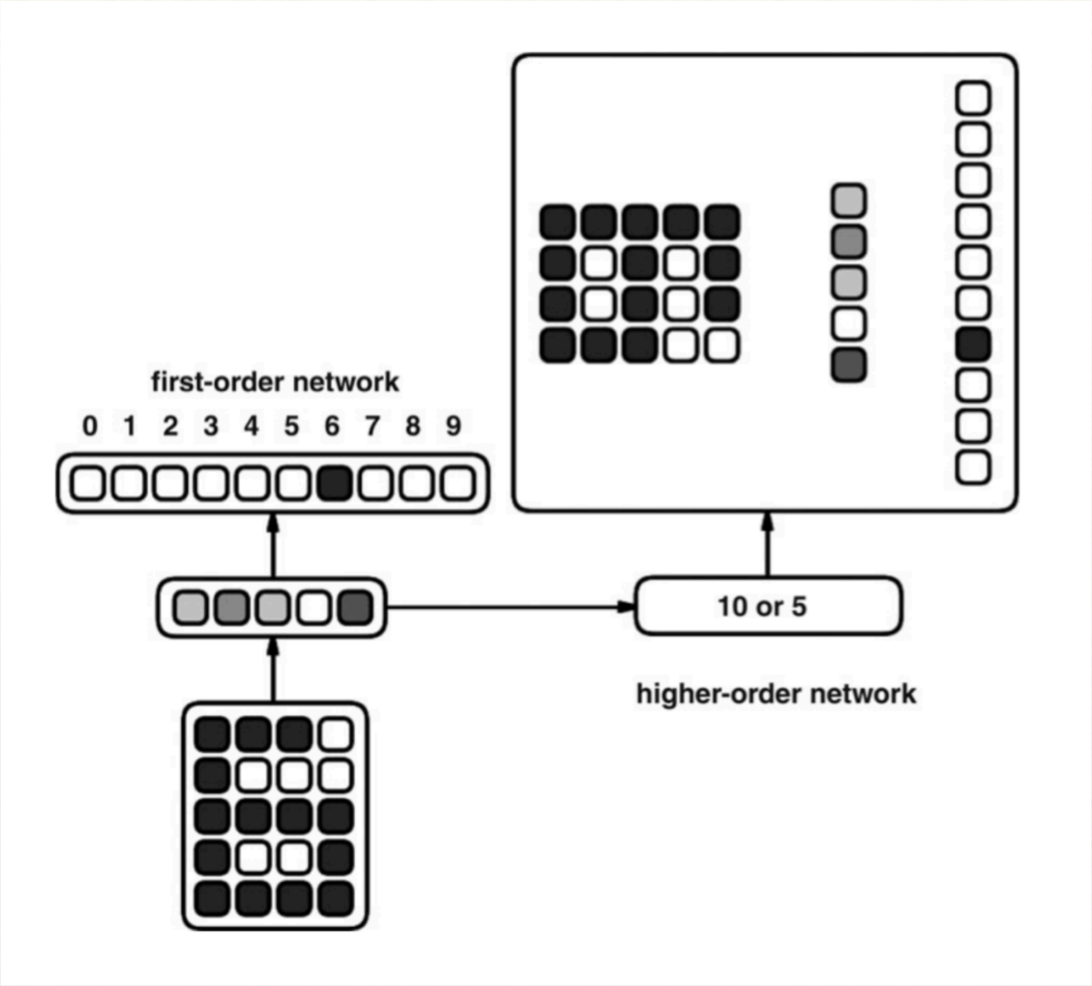
Part 3

[2] **Metarepresentation**: Machine needs to **know more about its own representation method**, namely to have a rough idea about what process it has undergone.

① Given one single state, conveying more information to the machine is conducive to **building the associations between experience**(input; what people see) **and memories**(outputs; what people feel like, might be a lot).

② Based on the understanding of what its own representations stand for, a machine can **access, inspect and manipulate** its representations. Finally, the machine might be able to form its own taste, thus rendering **unique outputs** with a given input. It will generate a 'self'.

Part 3



[2] **Figure:** Learn a **higher-order network** to take the **hidden units** of a first-order network as the input units. Its output is **all the units** in the first-order network, which means the output needs to simulate the whole process of the representation.

Part 4

What are the difficulties and challenges realizing autonomous learning?

Part 4

- [3] To evaluate a **continuously** learning model, we need to ensure that it works for:

  ① **Constant Memory**: As we store the knowledge, the memory should not **overflow**.

  ② **Problem Agnostic**: Is the method robust enough or is it restricted to some specific tasks?

  ③ **On Pre-trained:** Given a pre-trained model, autonomous learning can utilize and modify it to adapt to a new sequence of tasks.

  ④ **Unlabelled Data**: No need for supervision.

  ⑤ **Adaptive**: Selectively learn and spare some capacity for adaption at the same time.

Part 4

- [4] Challenges when implementing **curiosity**:

  ① Learn efficiently a large repertoire of diverse skills from the **vast learning space**.

  ② Avoid being trapped in **unlearnable situations**, for instance, predicting the details of white noise pattern on a television screen.

  ③ How to balance between **exploration** and **exploitation**, namely the balance between long-term and short-term benefits. [5]

Part 4

- [11] Challenges when storing memories, especially when confronting small sample learning:

  ① How to determine a proper family of models?

  ② How to define, represent, and embed knowledge into a model?

  ③ How to design metric learning methods more suitable for SSL?

[1] Stanislas Dehaene, Hakwan Lau, and Sid Kouider. "What is consciousness, and could machines have it?". In: *Science* 358.6362 (2017), pp. 486–492.

[2] Cleeremans, A., Timmermans, B., & Pasquali, A. "Consciousness and metarepresentation: A computational sketch". In: *Neural Networks*, 20(9), pp. 1032–1039.

[3] Rahaf Aljundi, Francesca Babiloni, Mohamed Elhoseiny, Marcus Rohrbach, and Tinne Tuytelaars. "Memory aware synapses: Learning what (not) to forget." In: *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 139–154, 2018.

[4] Jacqueline Gottlieb, Pierre-Yves Oudeyer, Manuel Lopes, and Adrien Baranes. "Information-seeking, curiosity, and attention: computational and neural mechanisms". In: *Trends in Cognitive Science*, 17(11), pp. 585–593, 2013.

[5] Rein Houthooft, Xi Chen, Yan Duan, John Schulman, Filip De Turck, Pieter Abbeel. "Curiosity-driven Exploration in Deep Reinforcement Learning via Bayesian Neural Networks". In: *CoRR*, abs/1605.09674, 2016.

[6] C. Fernando, D. Banarse, C. Blundell, Y. Zwols, D. Ha, A. A. Rusu, A. Pritzel, and D. Wierstra. "Pathnet: Evolution channels gradient descent in super neural networks". In: *CoRR*, abs/1701.08734, 2017.

[7] Sung, Flood & Yang, Yongxin & Zhang, Li & Xiang, Tao & H. S. Torr, Philip & Hospedales, Timothy. "Learning to Compare: Relation Network for Few-Shot Learning". In: *CVPR*, 2018.

[8] Pablo Hernandez-Leal, Bilal Kartal, and Matthew E Taylor. "Is multiagent deep reinforcement learning the answer or the question? A brief survey". In: *arXiv preprint* arXiv:1810.05587 (2018).

[9] M. Schuster, K. K. Paliwal. "Bidirectional recurrent neural networks". In: *IEEE Transactions on Signal Processing*, 45 (11) (1997) 2673–2681.

[10] P. Peng, Q. Yuan, Y. Wen, Y. Yang, Z. Tang, H. Long, J. Wang. "Multiagent Bidirectionally-Coordinated Nets for Learning to Play StarCraft Combat Games". In: *arXiv preprint* arXiv:1703.10069, 2017.

[11] Jun Shu, Zongben Xu, and Deyu Meng. "Small sample learning in big data era". In: *arXiv preprint* arXiv:1808.04572, 2018.