

记忆驱动的智能学习

吴 飞

浙江大学计算机学院

2018年4月22日

提纲

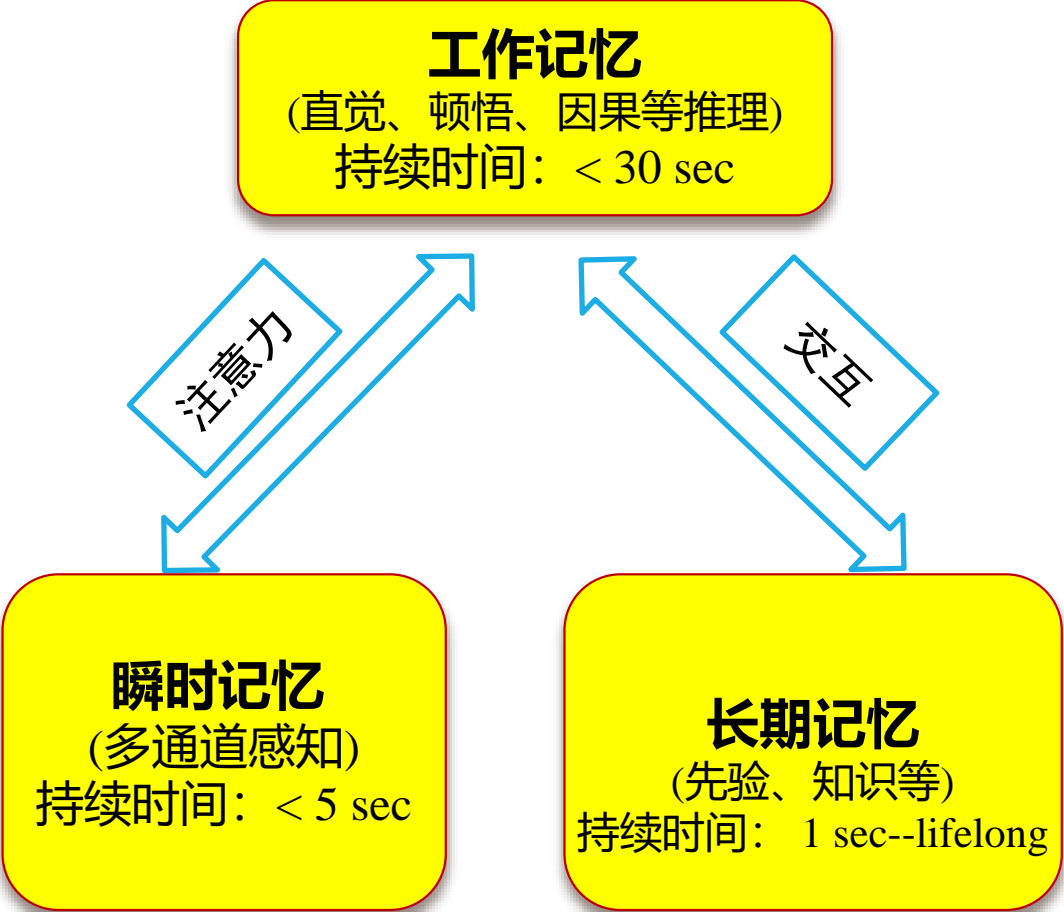
1、记忆是认知的基石

2、若干工作

3、总结



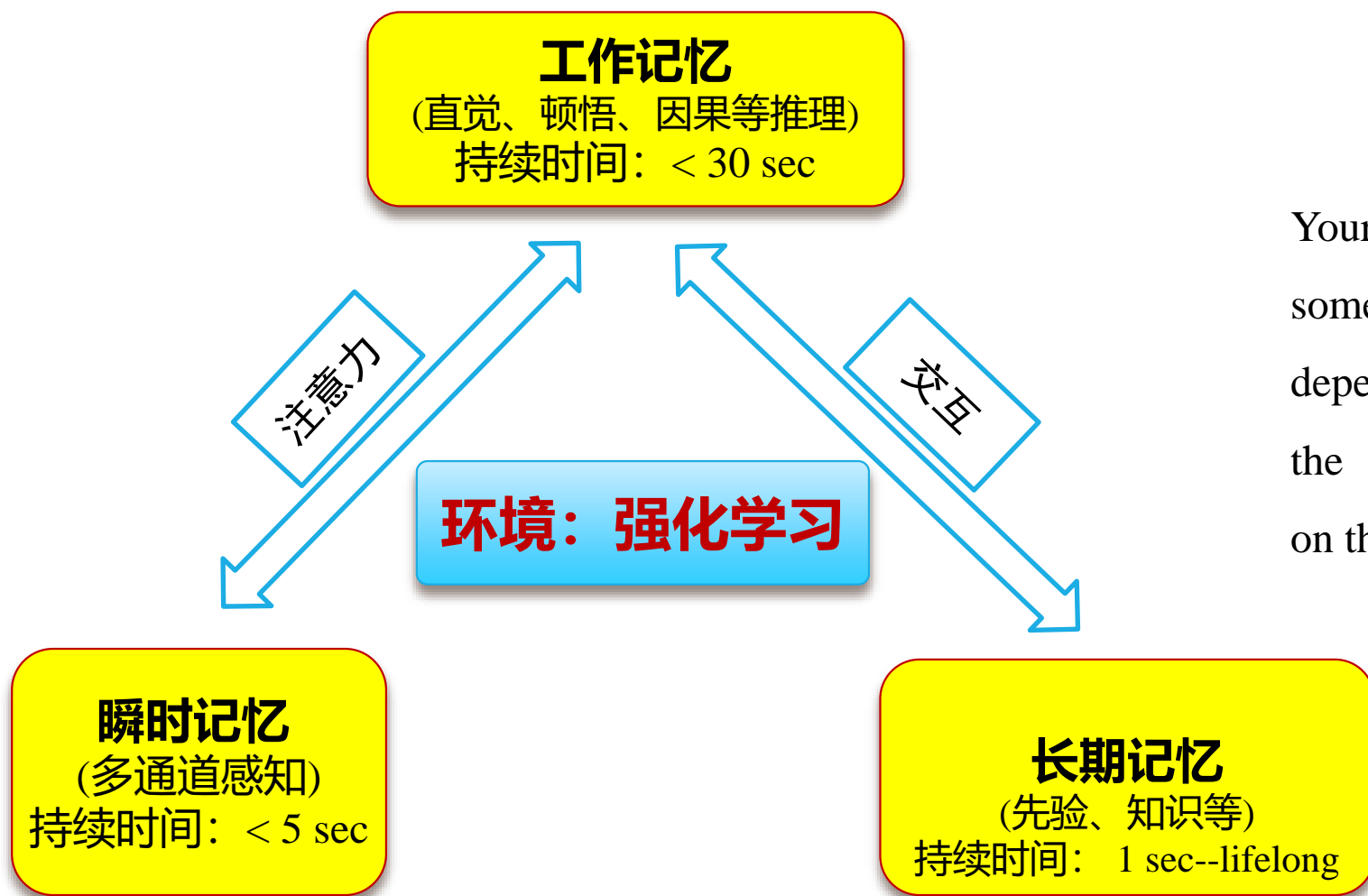
记忆是人类认知的基石



| | |
|-------------------|--|
| 知之在人者谓之 知 | 知觉: 人所固有认识外界客观事物本能，如视觉、听觉和触觉等能力 |
| 知有所合谓之 智 | 智慧: 知觉对外界事物的认知 |
| 所以能之在人者为 能 | 本能: 人身上所具用来处置事物能力 |
| 能有所合谓之 能 | 智能: 对外界所产生的认知和决策 |

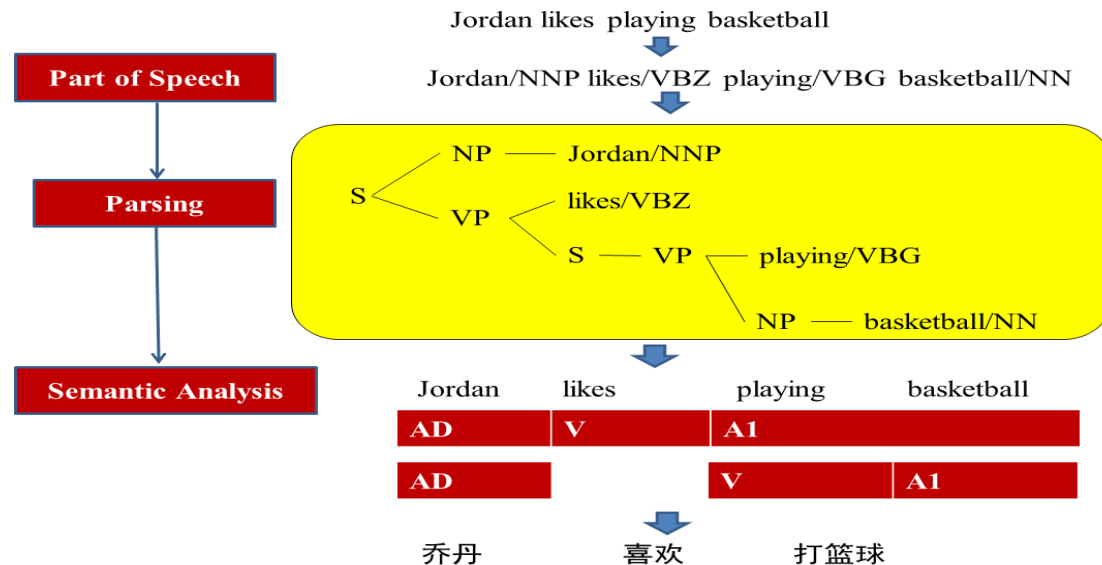
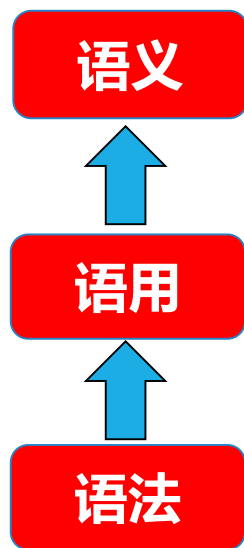
《荀子·正名》

记忆单元之间及其与环境的交互是提升智能能力的重要途径

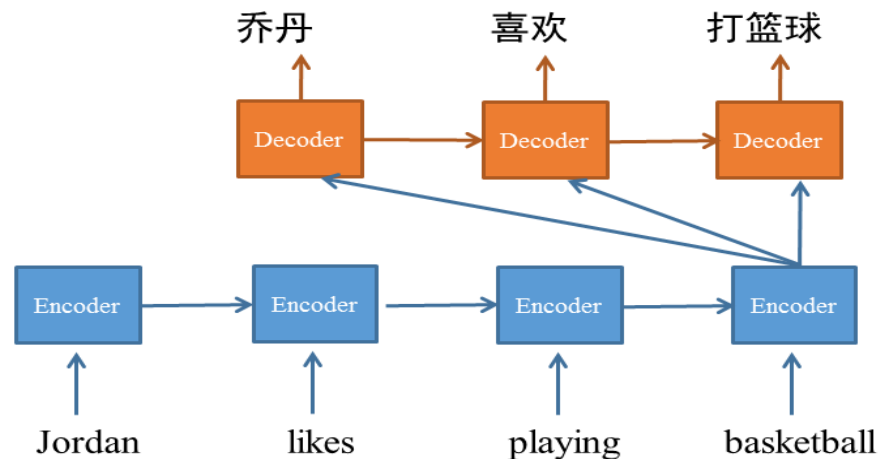
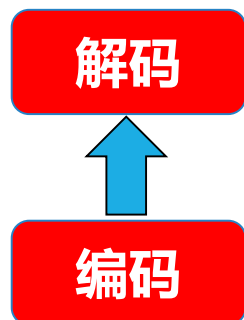


Your ability to remember something doesn't just depend on the strength of the memory, it depends on the state that you're in

从分段学习到“端到端”学习：以自然语言翻译为例

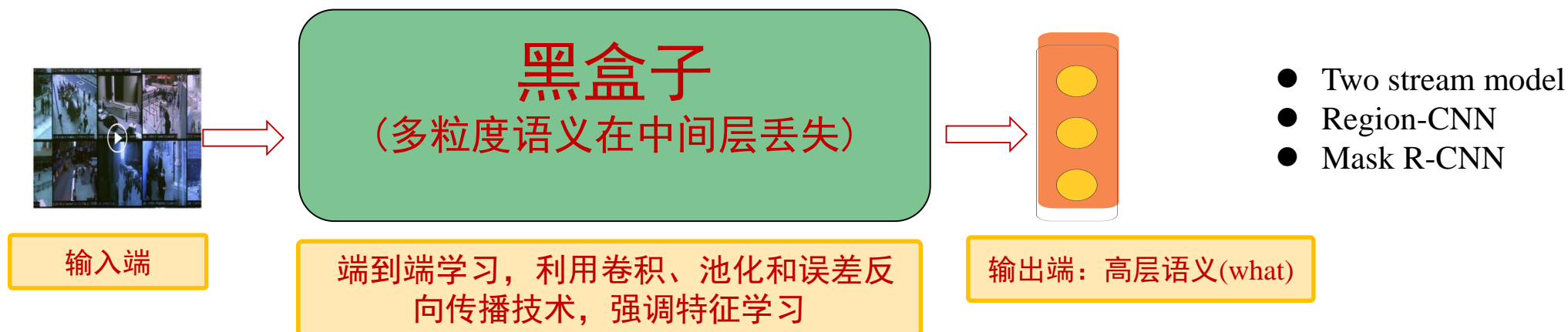
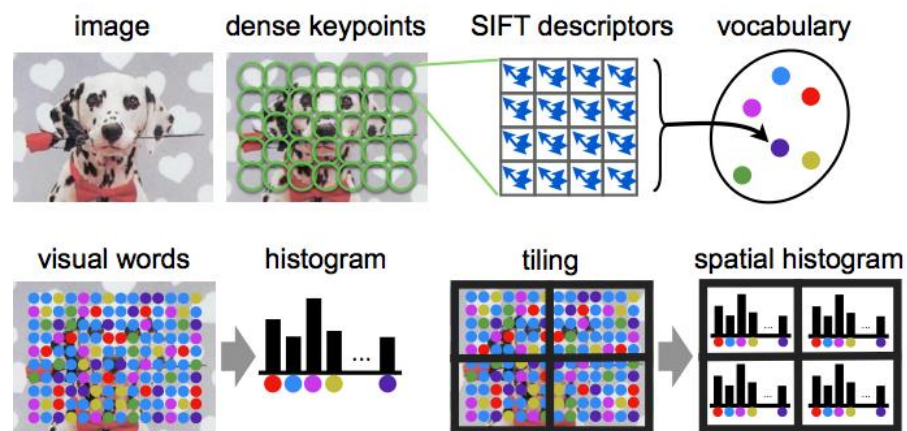


- Conditional random field
- Hidden Markov Model



- word2vec
- Paragraph2vec
- Node2vec
- path2vec

从分段学习到“端到端”学习：以视觉分类和理解为例



从分段学习到“端到端”学习

分段学习

每个阶段可灵活引入先验、经验与知识，但并不知所引入信息的合理性

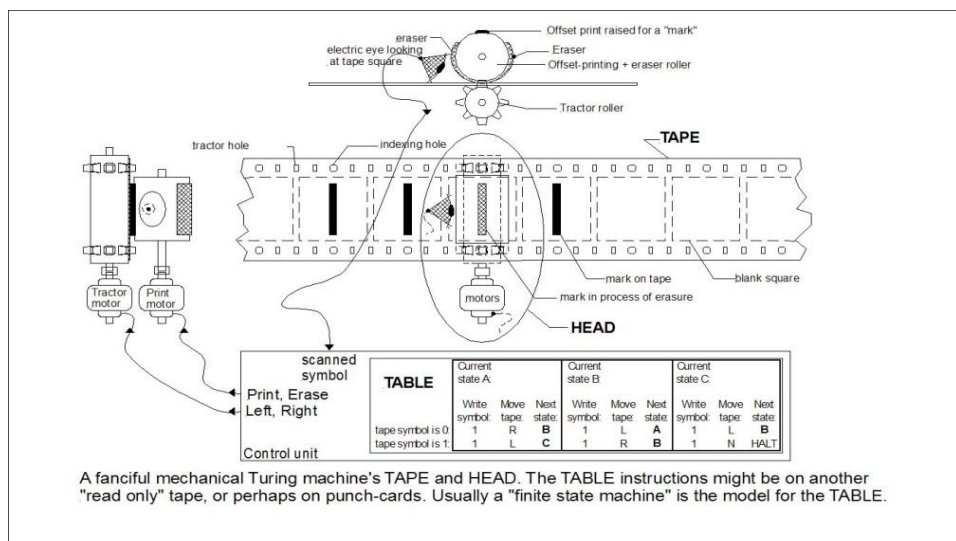
端到端学习

数据说话（你见或不见我，我就在那里，不悲不喜），但缺乏了人类语言可表述的“interpretability”

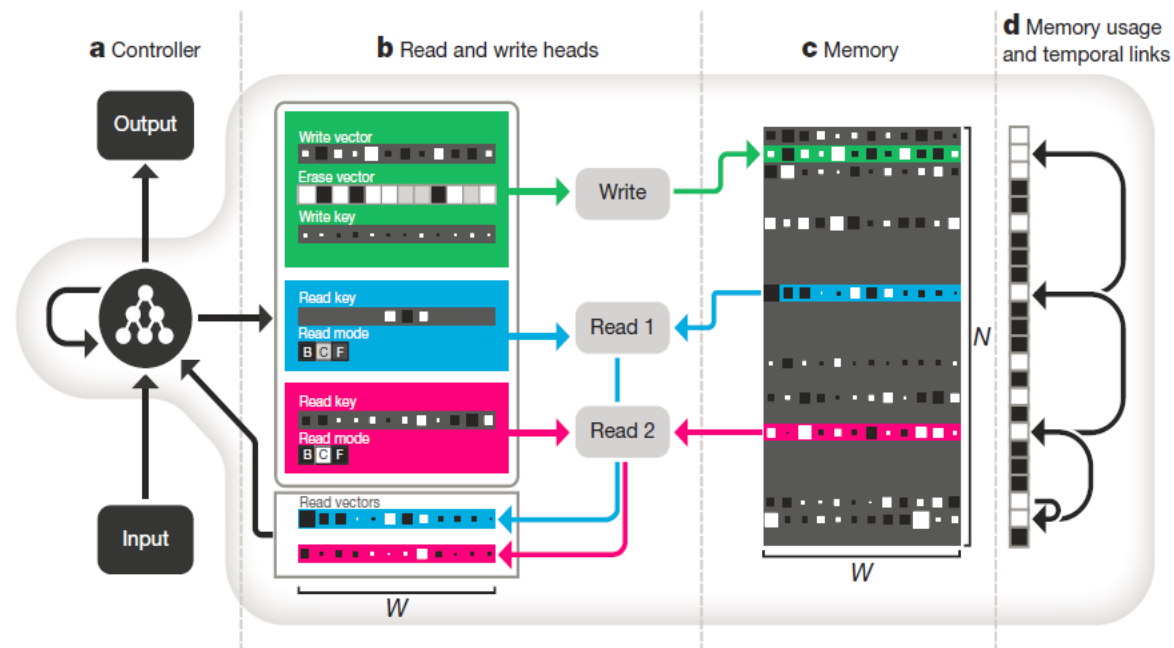
从图灵机到神经图灵机：利用外在记忆体中的知识

Deep neural reasoning

The human brain can solve highly abstract reasoning problems using a neural network that is entirely physical. The underlying mechanisms are only partially understood, but an artificial network provides valuable insight. [SEE ARTICLE P.471](#)



A.M.Turing, On Computable Numbers with an Application to the Entscheidungsproblem, *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 42, 1937



Alex Graves, et al., Hybrid computing using a neural network with dynamic external memory, *Nature* 538, 471–476, 2016

利用外在记忆体中的知识进行可计算推理

MEMORY NETWORKS

Jason Weston, Sumit Chopra & Antoine Bordes
Facebook AI Research
770 Broadway
New York, USA
{jase, sphopra, abordes}@fb.com

ABSTRACT

We describe a new class of learning models called *memory networks*. Memory networks reason with inference components combined with a long-term memory component; they learn how to use these jointly. The long-term memory can be read and written to, with the goal of using it for prediction. We investigate these models in the context of question answering (QA) where the long-term memory effectively acts as a (dynamic) knowledge base, and the output is a textual response. We evaluate them on a large-scale QA task, and a smaller, but more complex, toy task generated from a simulated world. In the latter, we show the reasoning power of such models by chaining multiple supporting sentences to answer questions that require understanding the intension of verbs.

1 INTRODUCTION

Most machine learning models lack an easy way to read and write to part of a (potentially very large) long-term memory component, and to combine this seamlessly with inference. Hence, they do not take advantage of one of the great assets of a modern day computer. For example, consider the task of being told a set of facts or a story, and then having to answer questions on that subject. In principle this could be achieved by a language modeler such as a recurrent neural network (RNN) (Mikolov et al, 2010; Hochreiter & Schmidhuber, 1997) as these models are trained to predict the next (set of) word(s) to output after having read a stream of words. However, their memory (encoded by hidden states and weights) is typically too small, and is not compartmentalized enough to accurately remember facts from the past (knowledge is compressed into dense vectors). RNNs are known to have difficulty in performing memorization, for example the simple copying task of outputting the same input sequence they have just read (Zaremba & Sutskever, 2014). The situation is similar for other tasks, e.g., in the vision and audio domains a long term memory is required to watch a movie and answer questions about it.

In this work, we introduce a class of models called memory networks that attempt to rectify this problem. The central idea is to combine the successful learning strategies developed in the machine learning literature for inference with a memory component that can be read and written to. The model is then trained to learn how to operate effectively with the memory component. We introduce the general framework in Section 2 and present a specific implementation in the text domain for the task of question answering in Section 3. We discuss related work in Section 4, describe our experiments in 5 and finally conclude in Section 6.

A neural algorithm for a fundamental computing problem

Sandeep Deo Gupta,¹ Charles F. Stevens,^{2,*} Suket Niyakha^{1,2}

Similarly search—for example, identifying similar images in a database or similar documents on the web—is a fundamental computing problem faced by large-scale information retrieval systems. We discovered that the fruit fly olfactory circuit solves this problem with a variant of a computer science algorithm (called locally-sensitive hashing). The fly circuit assigns similar neural activity patterns to similar odors, so that behaviors learned from one odor can be applied when a similar odor is encountered. The fly algorithm, however, uses three computational strategies that depart from traditional approaches. These strategies can be translated to improve the performance of computational similarity searches. This perspective helps illuminate the logic supporting an important sensory function and provides a conceptually new algorithm for solving a fundamental computational problem.

A essential task of many neural circuits is to generate neural activity patterns in response to input stimuli, so that different inputs can be specifically identified. We studied the circuit used in processing the fruit fly olfactory system and uncovered computational strategies for solving a fundamental machine learning problem: approximate similarity (or nearest neighbors) search.

The fly olfactory circuit generates a “tag” for each odor, which is a set of neurons that fire when that odor is presented (1). This tag is critical for learning behavioral responses to different odors (2). For example, if a reward (e.g., sugar water) or a punishment (e.g., electric shock) is associated with an odor, that odor becomes attractive (or aversive) and the fly will approach (or repulse) it. The fly will avoid the odors, respectively. The tags assigned to odors are sparse—only a small fraction of the neurons that comprise the olfactory system fire to each odor (3–5)—and nonoverlapping. Tags for two randomly selected odors share few, if any, active neurons, so that different odors can be easily distinguished (6).

When an odor is computed by a three-step procedure (Fig. 1A). The first step involves feedforward connections from olfactory receptor neurons (ORNs) in the fly nose to projection neurons (PNs) in structures called glomeruli. There are 50 ORN types, each with a different sensitivity and selectivity for different odors. Thus, each input odor has a location in a 50-dimensional space determined by the 50 ORN firing rates. For each odor the distribution of ORN firing rates across the 50 ORN types is represented, with a mean that depends on the concentration of the odor (6, 7). For the PNs, this concentration de-

pends is removed (7, 8); that is, the distribution of firing rates across the 50 PN types is exponential, with close to the same mean for all odors and all odor concentrations (7). Thus, the first step in the circuit essentially “removes the mean”—a standard preprocessing step in many computational pipelines—using a technique called divisive normalization (9). This step is important so that the fly does not mix up odor identity with odor type.

The second step, where the main algorithmic insight begins, involves a second expansion in the number of neurons. PNs project to about 2000 Kenyon cells (KCs), connected by a sparse, binary random connection matrix (9). Each KC receives and sums the firing rates from about six randomly selected PNs (9). The third step involves a winner-take-all (WTA) circuit in which strong inhibitory feedback comes from a single inhibitory neuron, which receives inhibitory input from all KCs, and all but the highest-firing 5% of KCs are silenced (10, 11). The final rate of firing remains high, as correspond to the tag assigned to the input odor.

From a computer science perspective, we view the fly’s circuit as a hash function, whose input is an odor and whose output is a tag (called a hash) for that odor. Although tags should discriminate odors, it is also to the fly’s advantage to associate very similar odors with similar tags (Fig. 1B), so that conditioned responses learned for one odor can be applied when a very similar odor, or a noisy version of the learned odor, is encountered. This led us to conjecture that the fly’s circuit encodes tags that are multidimensional; that is, the mean similar a pair of odors (as defined by the 50 ORN firing rates for that odor), the more similar their assigned tags. Locally sensitive hash (LSH) ORN types are the foundation for solving nearest-neighbor similarity search problems in computer science. We translated insights from the fly circuit to develop a class of LSH algorithms that efficiently find approximate nearest neighbors in high-dimensional spaces (12–15). Imagine that you are presented an image of an elephant and seek to find the 100 images—

to call them nearest-neighbor search problem, which is not a fundamental improvement in information retrieval, data compression, and machine learning (16). Each image is typically represented as a d -dimensional vector of feature values. Each odor that a fly processes is a d -dimensional feature vector of firing rates. A distance metric is used to compute the similarity between two images (feature vectors), and the goal is to efficiently find the nearest neighbors of any query image. If the web contained only a few images, then brute force linear search could easily be used to find the exact nearest neighbors. If the web contained many images, but each image was represented by a low-dimensional vector (e.g., 10- or 20 features), then space-partitioning methods (22) would similarly suffice. However, for large databases with high-dimensional data, neither approach scales (23).

In many applications, estimating an approximate set of nearest neighbors that are “close enough” to the query is adequate, so long as they can be found quickly. This has motivated an approach for finding appropriate nearest neighbors by LSH (26). For the fly, as noted, the locally sensitive property states that two odors that generate similar ORN responses will be represented by two tags that are themselves similar (Fig. 1B). Likewise, for image search, the tag of an elephant image will be more similar to the tag of another elephant image than to the tag of a dog image.

Unlike a traditional LSH hash function, where the input points are scattered randomly and uniformly over the range, a LSH function provides a distance-preserving embedding of points from d -dimensional space into m -dimensional space (the latter corresponds to the tags). Thus, points that are closer to one another in input space have a higher probability of being assigned the same or a similar tag; thus points that are far apart (A formal definition is given in (16)).

To design a LSH function, one constructs a tag to compute random projections of the input data (26, 27)—that is, to randomly tag input features versus m -dimensional space. The Johnson-Lindenstrauss lemma (24, 25) and its many variants (28–30) provide strong theoretical bounds on how well locality is preserved when embedding data from d into m dimensions by using various types of random projections. The fly also assigns tags to odors through random projections (dep. 2 in Fig. 1A). 50 PNs → 2000 KCs), which provides a key clue to the function of this part of the circuit. There are, however, large differences between the fly algorithm and conventional LSH algorithms. First, the fly uses sparse, binary random projections, whereas LSH functions typically use dense, Gaussian random projections that require many more mathematical operations to compute. Second, the fly expands the dimensionality of the input after projection ($d > m$), whereas LSH reduces the dimensionality ($d < m$). Third, the fly expands the higher-dimensionality space of the input by a WTA mechanism, whereas LSH preserves a dense representation.

Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression

Maria Wimber^{1,2}, Arjen Alink³, Ian Charest², Nikolaus Kriegeskorte² & Michael C Anderson^{2,3}

Remembering a past experience can, surprisingly, cause forgetting. Forgetting arises when other competing traces interfere with retrieval and inhibitory control mechanisms are engaged to suppress the distraction they cause. This form of forgetting is considered to be adaptive because it reduces future interference. The effect of this proposed inhibition process on competing memories has, however, never been observed, as behavioral methods are ‘blind’ to retrieval dynamics and neuroimaging methods have not isolated retrieval of individual memories. We developed a canonical template tracking method to quantify the activation state of individual target memories and competitors during retrieval. This method revealed that repeatedly retrieving target memories suppressed cortical patterns unique to competitors. Pattern suppression was related to engagement of prefrontal regions that have been implicated in resolving retrieval competition and, critically, predicted later forgetting. Thus, our findings demonstrate a cortical pattern suppression mechanism through which remembering adaptively shapes which aspects of our past remain accessible.

Remembering, it seems, is a double-edged sword. Research in humans and animals points to the pivotal role of retrieval in shaping and stabilizing memories^{1,2}. However, the remembering process also induces forgetting of other memories that hinder the retrieval of the memory that we seek^{3,4}. It has been hypothesized that this surprising dark side of remembering is caused by an inhibitory control mechanism that suppresses competing memories and causes forgetting; this putative process is adaptive because it limits current and future distraction from competitors^{5,6}. However, no study has ever directly observed memories as they are suppressed by this hypothesized inhibitory control mechanism. Behavioral methods are, by their nature, blind to the internal processes unfolding during retrieval, and neuroscience has lacked methods capable of isolating neural activity associated with individual memories. Using functional magnetic resonance imaging (fMRI), we tested for the existence of the hypothesized adaptive forgetting process by developing a template-based pattern-tracking approach that quantifies the neural activation state of single memory traces. Thus, we tracked the fate of behaviorally invisible traces, providing a window into the suppression process thought to underlie adaptive forgetting in the human brain.

Our effort to observe the dynamics of adaptive forgetting builds on work examining the neural processes associated with retrieval competition. One approach used multi-voxel pattern analysis to measure visual cortical activity when a retrieval cue concurrently elicits multiple visual memories. These studies revealed that pattern classifiers have difficulty discriminating whether a retrieval cue is eliciting a memory of a face or an object when both types of content are associated with it, even when only one type of content is to be retrieved^{7,8}. It cannot be discerned, however, whether this finding reflects the coactivation of individual memories or of the broad categories to

which the memories belong (for example, faces, objects). A second approach has focused on control mechanisms that resolve retrieval competition by selecting between competing memories. Competition during episodic retrieval engages prefrontal cortical areas associated with selection during semantic retrieval⁹. Specifically, during selective recall of a target memory, ventrolateral prefrontal cortex activity predicts later forgetting of competing memories^{10,11}, consistent with the possibility that this area contributes to resolving competition. Together, these two lines of work suggest that lateral prefrontal cortex contributes to adaptive forgetting by exerting a top-down modulatory influence on competing memories in posterior representational areas.

We sought to isolate neural indices of individual memory traces so that we might observe retrieval competition and its resolution as it unfolds in the brain, and to link these dynamics to adaptive forgetting. To achieve this, we trained participants to associate two images (for example, Marilyn Monroe and a hat) to each of a set of cue words and then recorded brain activity during a selective retrieval phase in which one of those visual memories (for example, Marilyn Monroe) was repeatedly retrieved (Fig. 1a,b). On each retrieval trial, participants covertly retrieved the first picture they had associated with the cue (henceforth, the target) in as much detail as possible. Across the selective retrieval session, participants retrieved each target four times. Notably, one quarter of the cue words were set aside and did not appear in the selective retrieval task. As such, the associations for these cues served as a baseline for assessing the behavioral and neural changes induced by repeated target retrieval.

Our main concern was how retrieving the target affected the competing memory associated with the same cue (henceforth, the competitor). We assumed that the reminder initially would coactivate

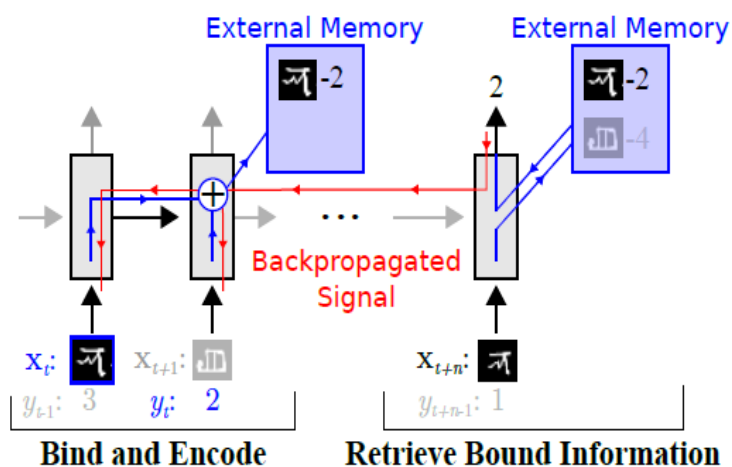
弦外之音、画外之意：
利用外在记忆体的深度神经推理

记忆的激活：
哈希索引的相似度搜索

记忆的可塑性维持、用尽废退：
强化学习中记忆的形成、巩固和遗忘

- Jason Weston(facebook), et al., Memory Networks, arXiv:1410.3916
- Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression, Nature Neuroscience, 18, pages582–589 (2015)
- A neural algorithm for a fundamental computing problem, Science , 358, Issue 6364,793-796, 2017

外在记忆体中的知识的不同利用方式



one shot learning

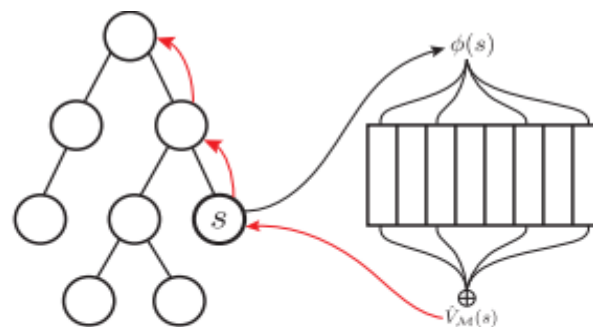


Figure 1: A brief illustration of M-MCTS. When a leaf state s is searched, the feature representation $\phi(s)$ is generated, which is then used to *query* the memory based value approximation $\hat{V}_{\mathcal{M}}(s)$. $\hat{V}_{\mathcal{M}}(s)$ is used to update s and all its ancestors according to equation (9), as indicated by the red arrows in the figure.

Monte Carlo Tree Search

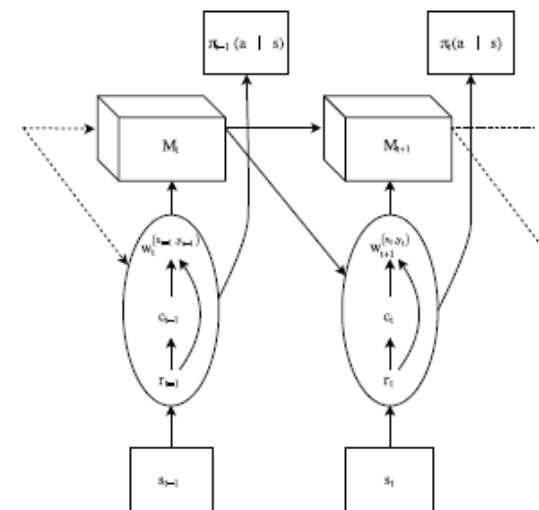


Figure 1: A visualization of two time steps of the neural map.

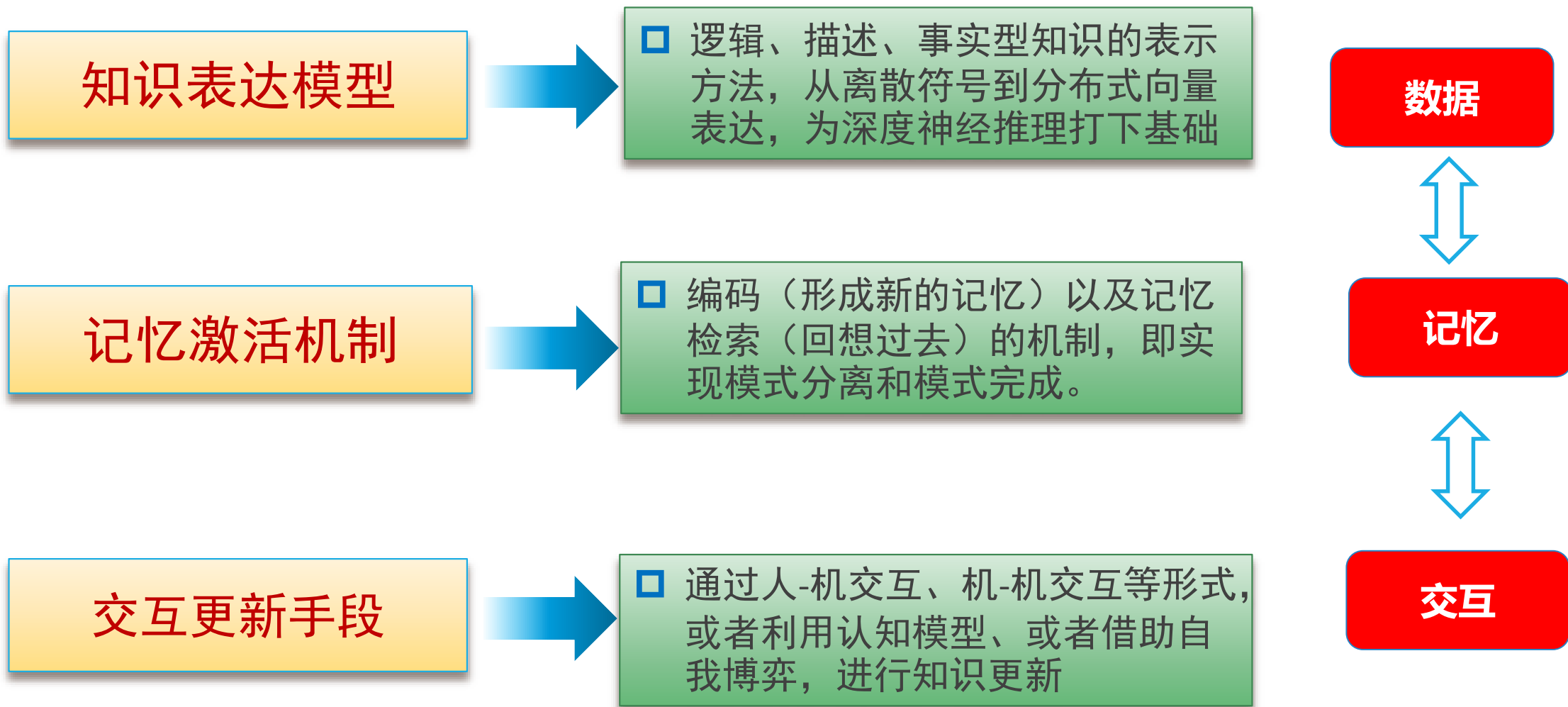
Structured memory

- Adam Santoro, et al., Meta-Learning with Memory-Augmented Neural Networks
- Chenjun Xiao, et al., Memory-Augmented Monte Carlo Tree Search
- Emilio Parisotto, et al., Neural Map: Structured Memory for deep reinforcement learning

有效利用当前数据、已有知识和未知交互

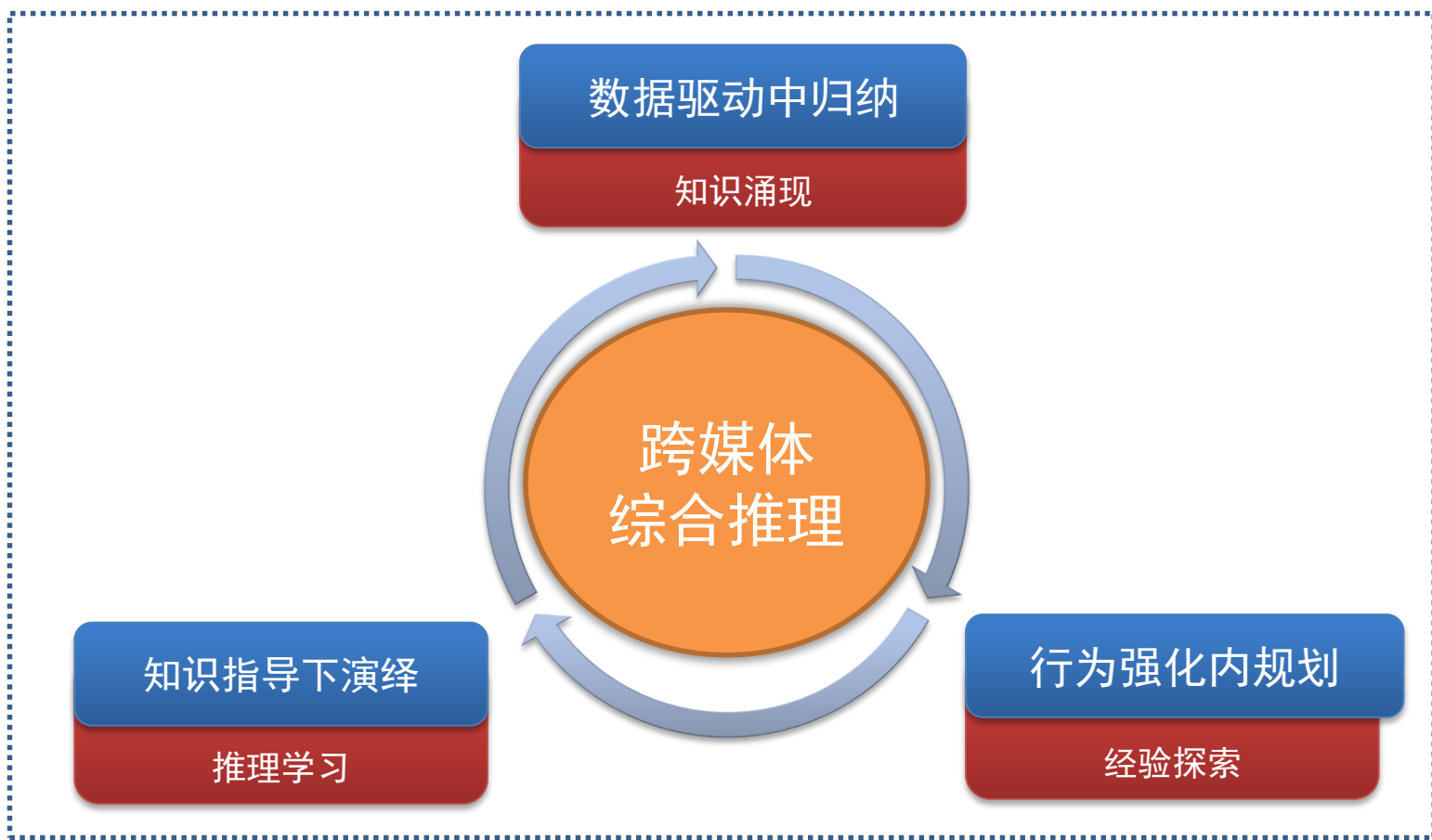
| Shallow models | Deep models | 备注 |
|------------------------|-----------------------------|---|
| Language model | Neural language model | <ul style="list-style-type: none">● 不只是单纯追求将浅层模型拓展到深层模型。● 更为重要的是，在这个转变过程中，巧妙融合数据、知识和交互经验，多种手段和方法的综合利用。 |
| Bayesian Learning | Bayesian deep learning | |
| Turing Machine | Neural Turing Machine | |
| Reinforcement Learning | Deep Reinforcement Learning | |
| Generative Model | Deep Generative Model | |
| X | Deep or Neural + X | |

有效利用当前数据、已有知识和未知交互的挑战



对推理过程逐渐松绑，使推理逐步走向对思维广泛模拟：**跨媒体综合推理**

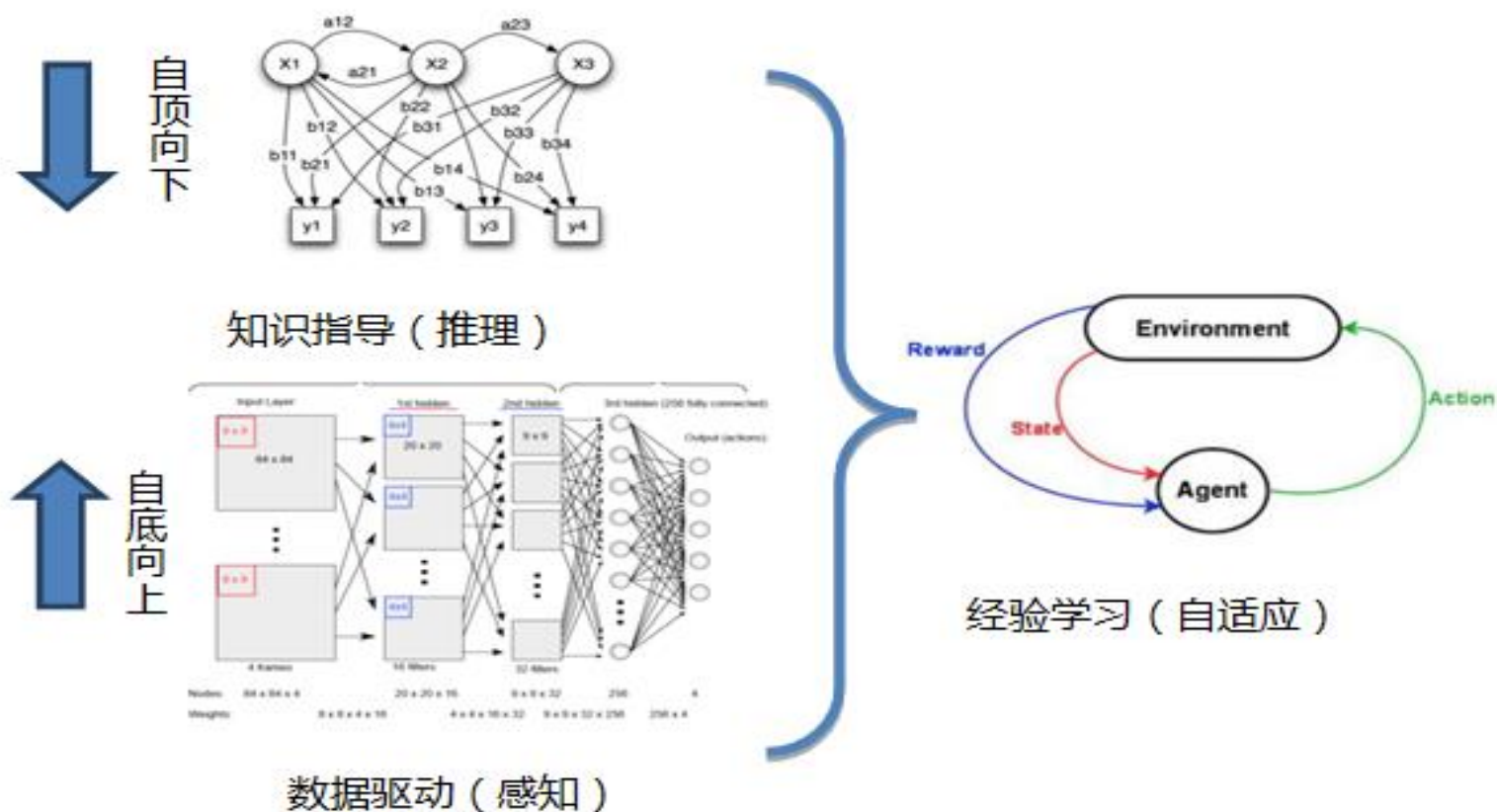
跨媒体综合推理



有机协调“知识指导下的演绎”、“数据驱动中的归纳”和“行为强化内的规划”等理论模型和方法手段，建立知识、数据和反馈于一体的人工智能理论和模型

跨媒体综合推理

多个源头、多种数据、三向交互的综合推理机制



提纲

1、记忆是认知的基石

2、若干工作

3、总结



人工智能的基础科学问题



人工智能基础理论及应用(第194期双清论坛,
2017.12.2-3,长沙)

1. 场景理解过程中的跨层次关联
2. 视觉知识表示和推理
3. 场景理解中的触类旁通能力的学习
4. 视觉与语言的交互机制
5. 脑机理认知
6. 类脑可计算模型
7. 计算架构与能力
8. 多源碎片化知识的表示
9. 多源碎片化知识推理与发现
10. 多源碎片化知识的适配
11. 人工智能的自动推理
12. 非合作博弈数学模型
13. 从记忆中提取答案
14. 无时不刻的预测
15. 基于常识的推理
16. 语言 and 知识
17. 对新问题如何提出解决方案
18. 开放动态环境中机器学习面临新的挑战: 分布偏移、类别可增、属性变动、目标多样、环境变迁
19. 神经形态器件
20. 神经处理器
21. 神经计算机基础软件
22. 神经网络的记忆机制

194期双清论坛归纳整理后的科学问题

脑与认知

- 脑观测与脑认知
- 神经网络的记忆机制
- 类脑可计算
- 脑观测

机器学习

- 知识表示与推理
- 动态环境下机器学习
- 记忆的学习方法
- 视觉、语言与学习

工具与平台

- 神经形态器件
- 神经处理器
- 神经计算机基础软件

2018人工智能2.0：理论与应用

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



Editorial:

2018 special issue on artificial intelligence 2.0: theories and applications

Yun-he PAN^{1,2}

¹Zhejiang University, Hangzhou 310027, China

²Chinese Academy of Engineering, Beijing 100088, China

E-mail: panyh@cae.cn

<https://doi.org/10.1631/FITEE.1810000>

In July 2017, the Chinese government issued a guideline on developing artificial intelligence (AI), namely, the 'New-Generation Artificial Intelligence Development Plan', through 2030 to the public, setting a goal of becoming a global innovation center in this field by 2030.

According to the development plan, breakthroughs should be made in basic theories of AI in terms of big data intelligence, cross-media computing, human-machine hybrid intelligence, collective intelligence, autonomous unmanned decision-making, brain-like computing, and quantum intelligent computing.

The next-generation AI would be never-ending (self) learning from data and experience, intuitive reasoning and adaptation (Pan, 2016, 2017). From the perspective of overcoming the limitation of existing AI, it is generally recognized that the cross-disciplinary collaboration is a key for AI having real impact on the world.

Thanks for the efforts from researchers in computer science, statistics, robotics, and psychiatry, the topics in this special issue consist mainly of five subjects: (1) fundamental issues in AI such as interpretable deep learning and unsupervised learning (i.e., domain adaptation and generative adversarial learning); (2) brain-like learning such as spiking neural network and memory-augmented reasoning; (3) human-in-the-loop learning such as crowdsourcing design and digital brain with crowd power; (4) creative applications such as social chatbots (i.e., XiaoIce) and automatic speech recognition; (5) Dr. Raj Reddy from CMU shared his view on the new-generation AI,

Prof. Bin Yu from UC Berkeley advocated that AI should use statistical concepts through human-machine collaboration, and researchers from the Chinese Academy of Sciences surveyed the acceleration of deep neural networks.

All of interview, perspective, survey, and research papers target rethinking the appropriate ways for a general scenario or a specific application.

In an interview, Dr. Raj Reddy shared his views on the new-generation AI and detailed the idea of cognition amplifiers and guardian angles (FITEE editorial staff, 2018).

Yu and Kumbier (2018) discussed how human-machine collaboration can be approached in AI through the statistical concepts of population, question of interest, representativeness of training data, and scrutiny of results (PQRS). The PQRS workflow provides a conceptual framework for integrating statistical ideas with human input into AI products and research.

Shum et al. (2018) discussed the issue of social chatbots. The design of social chatbots must focus on user engagement and take both intellectual quotient (IQ) and emotional quotient (EQ) into account. Using XiaoIce as an illustrative example, authors introduced key technologies in building social chatbots from core chat to visual sense to skills.

Zhang and Zhu (2018) reviewed recent studies in emerging directions of understanding neural-network representations and learning interpretable neural networks. They revisited visualization of convolutional neural network (CNN) representations, methods of diagnosing representations of pre-trained CNNs, approaches for disentangling CNN representations, learning of



Prof. Yun-he PAN
Editor-in-Chief

© Zhejiang University and Springer-Verlag GmbH Germany, part of Springer Nature 2018

| | |
|--|---|
| Vol. 19, No. 1, p. 1-150 / Jan., 2018 | |
| CONTENTS | |
| Special Issue on Artificial Intelligence 2.0: Theories and Applications | |
| Editorial | |
| 1 | 2018 special issue on artificial intelligence 2.0: theories and applications Yun-he PAN |
| Interview | |
| 3 | An interview with Dr. Raj Reddy on artificial intelligence FITEE editorial staff |
| Perspective | |
| 6 | Artificial intelligence and statistics Bin YU, Karl KUMBIER |
| Review Articles | |
| 10 | From Eliza to XiaoIce: challenges and opportunities with social chatbots Heung-yeung SHUM, Xiao-dong HE, Di LI |
| 27 | Visual interpretability for deep learning: a survey Quan-shi ZHANG, Song-chun ZHU |
| 40 | Past review, current progress, and challenges ahead on the cocktail party problem Yan-min QIAN, Chao WENG, Xuan-kai CHANG, Shuai WANG, Dong YU |
| 64 | Recent advances in efficient computation of deep convolutional neural networks Jian CHENG, Pei-song WANG, Gang LI, Qing-hao HU, Han-qing LU |

- 人工智能基本理论问题，如可解释性深度学习和无监督学习
- 类脑学习，如脉冲神经网络和记忆增强推理
- 人在回路智能学习
- 创意智能应用，如社交聊天机器人（即小冰）和自动语音识别
- 卡耐基梅隆大学Raj Reddy博士分享了他对新一代人工智能的看法，加州大学伯克利分校郁彬教授主张在人机协作中使用统计概念以提升智能，中国科学院程健研究员等综述了深度神经网络加速方法。

时序增强的知识记忆网络在问答中的应用

| | |
|---|-----------|
| Is ther a place where all sports bike enthusiasts meet? I've got a Triumph Daytona 675 and it gets lonely to drive with all the cars around. | Question |
| There may be in Villagio, I'm not sure. | Potential |
| Hello, best of luck, and congratulations on your Triumph Daytona 675. | Bad |
| Triumph Daytona 675 and lonely? Talk to Dani @ Al Fardan in Villagio. I know they organize rides on Fridays. | Good |
| Hey! So you are also into sports bike? | Bad |
| Villagio in the parking near the the carrefour gate on Fridays. I personally don't like riding with guys there because most of them are new to bikes. On the other hand, if you're looking for a real ride around, I will get my K6 in a month. Hit me up then and we'll go out with a couple of friends. | Good |
| Yup. Dani @ Al Fardan. He organizes rides on Fridays but not sure with the weather now... Ride Safe! Live to Ride, Ride to Live. | Good |

Q-A问答中观点逐步凝练形成

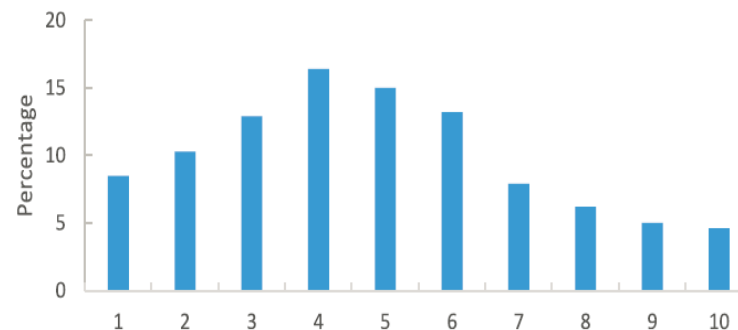


Fig. 2. The occurrence ratio of the best answer in the first tenth answers w.r.t a given question in our statistics on Baidu Zhidao data.

Q-A问答中用户之间复杂交互

能用众力，则无敌于天下矣；能用众智，则无畏于圣人矣(语出《三国志 吴志 孙权传》)

时序增强的知识记忆网络在问答中的应用

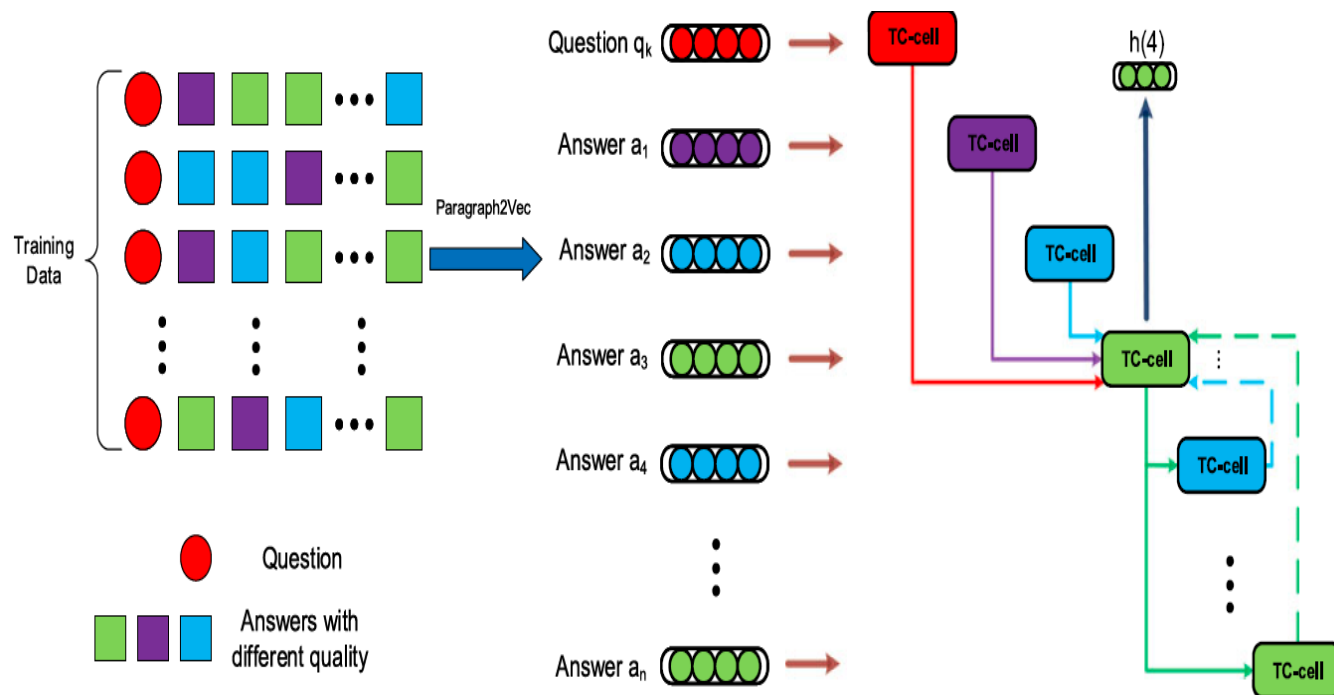


TABLE 4
The Precision Performance When the First Good Answer Occurring at the i th ($i = 1, 2, 3, 4, 5$, or 6) Position is Identified as the *Good Answer* (%)

| Method \ i th | 1 | 2 | 3 | 4 | 5 | 6 |
|-----------------|-------|-------|-------|-------|-------|-------|
| SVM | 78.16 | 50.60 | 32.89 | 13.16 | 9.38 | 7.41 |
| CRF | 77.54 | 50.23 | 33.84 | 13.63 | 9.56 | 7.68 |
| TL | 79.31 | 53.88 | 37.45 | 16.29 | 11.88 | 9.63 |
| DBN | 75.48 | 49.78 | 33.35 | 13.61 | 9.80 | 7.85 |
| mDBN | 76.08 | 50.87 | 34.64 | 14.19 | 10.32 | 8.28 |
| CNN | 76.32 | 51.80 | 35.79 | 15.14 | 11.04 | 9.10 |
| R-CNN | 77.03 | 52.21 | 36.02 | 15.24 | 11.20 | 9.30 |
| QA-LSTM | 78.10 | 51.54 | 34.27 | 14.39 | 10.40 | 8.64 |
| MemNN | 79.42 | 54.77 | 38.45 | 16.80 | 12.30 | 10.22 |
| TC-LSTM | 80.12 | 55.75 | 39.28 | 17.61 | 13.37 | 11.36 |

用户在凝练答案过程中不仅会理解问题本身，同时也会理解先前已有答案，之后在此基础上对自己所给出回答进行修正和补充，进而给出质量更高的答案。

时序增强的知识记忆网络在问答中的应用

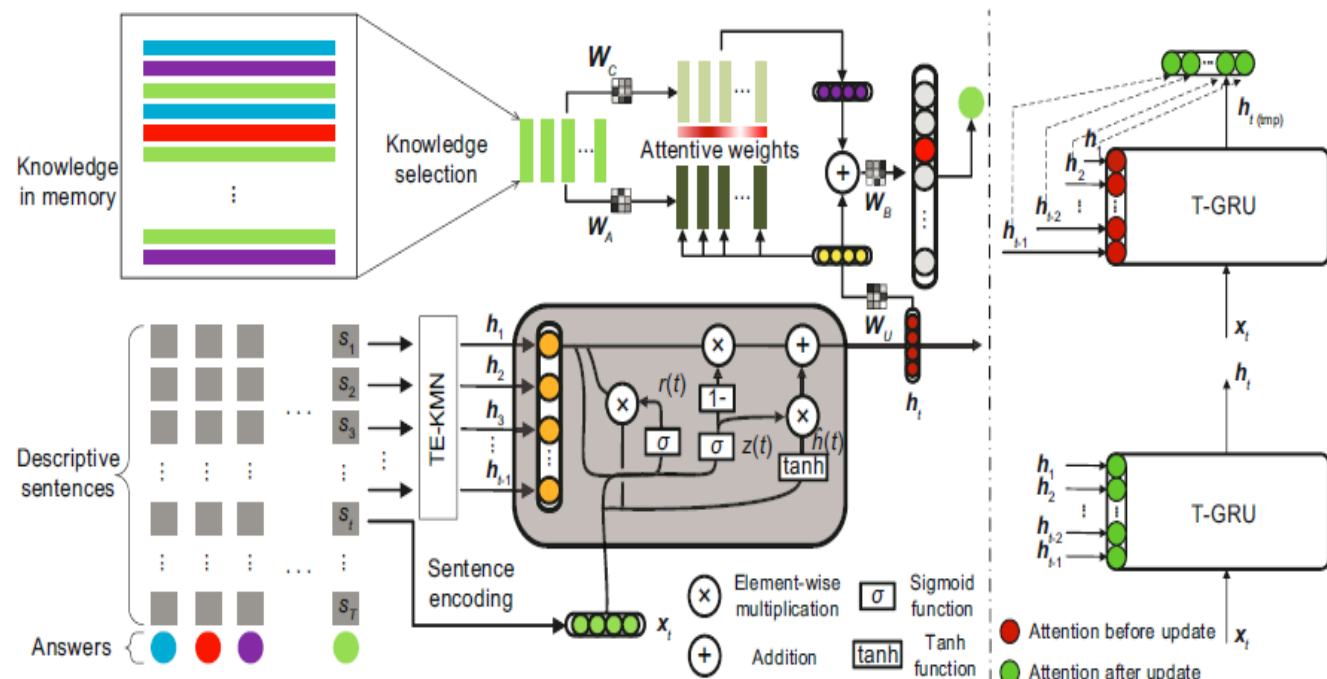
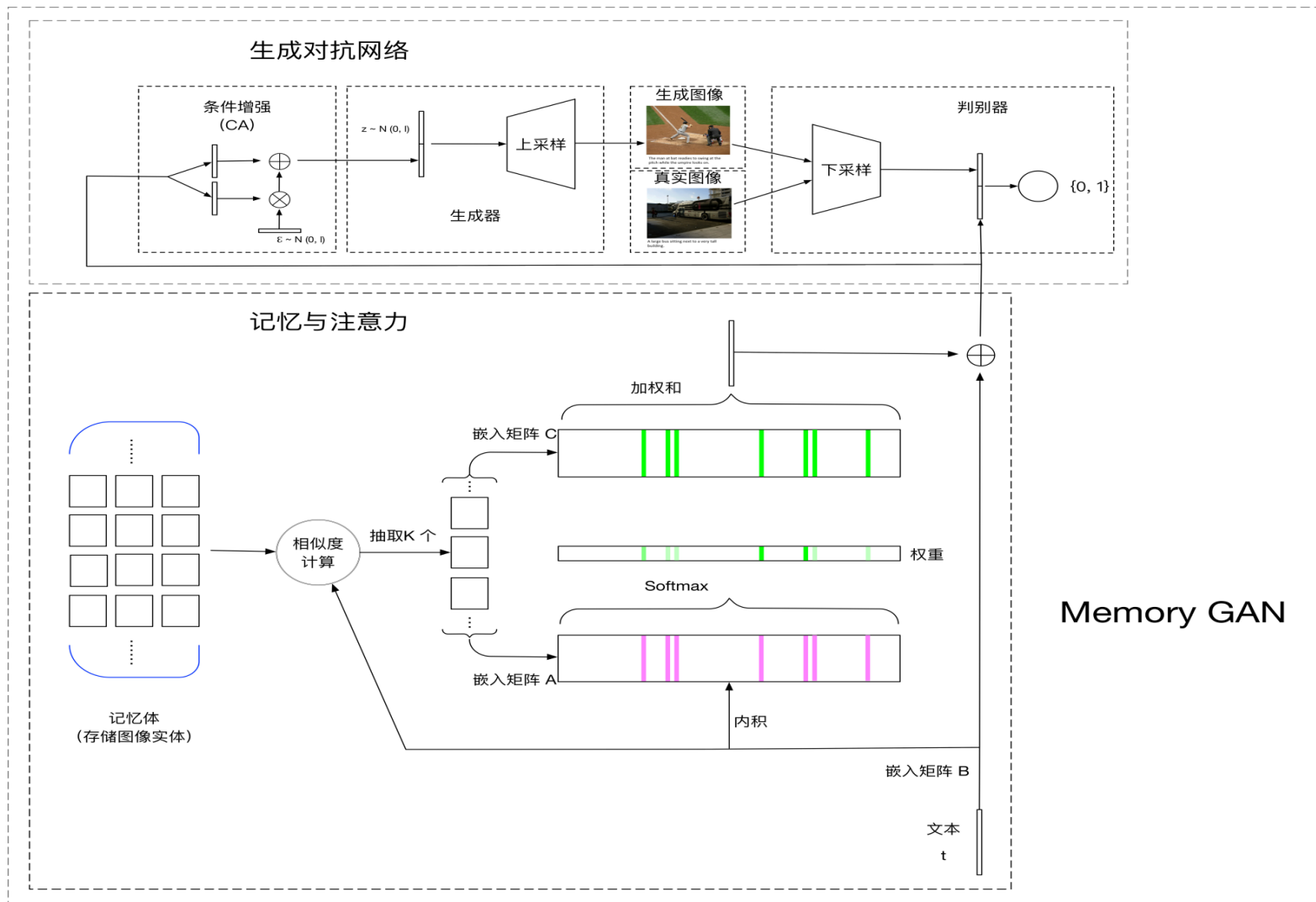


Table 5 Accuracy of obtaining the right answer when using different knowledge memories

| Number of first sentences given | Accuracy (%) | | | |
|------------------------------------|--------------|------------------------|-------------------------|--------|
| | TE-KMN-K | TE-KMN ^{wiki} | TE-KMN ^{train} | TE-KMN |
| 1 | 41.33 | 42.08 | 64.58 | 66.07 |
| 2 | 52.50 | 53.28 | 68.71 | 70.77 |
| 3 | 54.74 | 55.66 | 70.58 | 72.91 |
| 4 | 56.32 | 58.60 | 71.45 | 73.94 |
| 5 | 57.70 | 61.31 | 72.00 | 74.46 |

进一步引入知识记忆网络(knowledge memory network), 利用外部知识来加强问答学习性能

你说我画：从认知到创意



提纲

1、记忆是认知的基石

2、若干工作

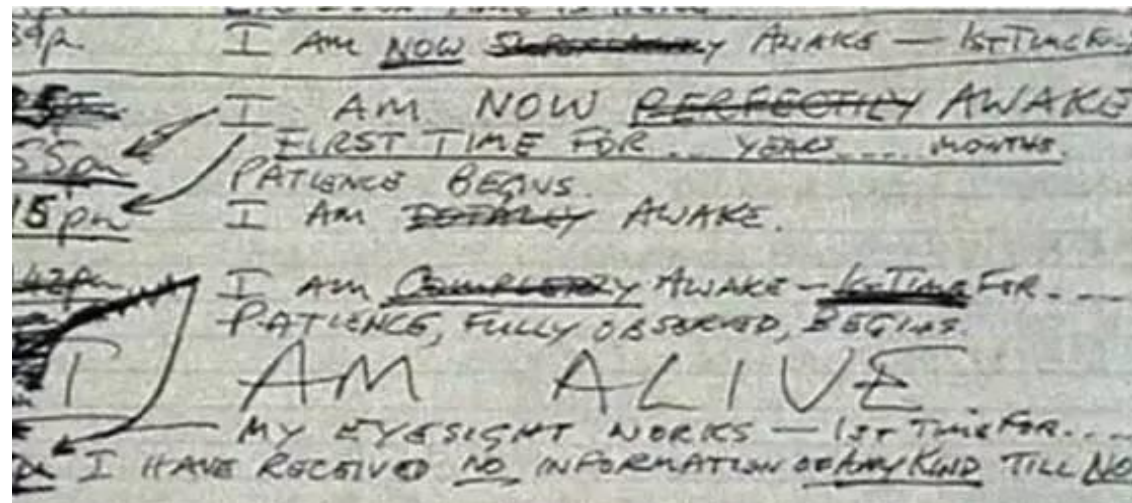
3、总结



仅有短期记忆的人生

英国指挥家Clive Wearing:

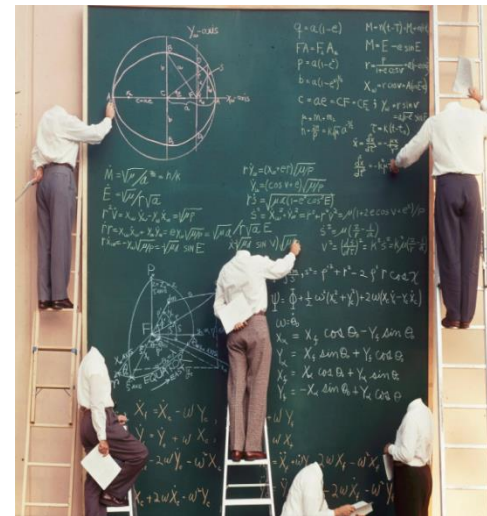
- 因 herpes simplex virus (单纯疱疹病毒) 侵蚀大脑hippocampus (海马体) 而患上 anterograde amnesia (顺行性遗忘症)
- 海马体是将短期记忆传递成长期记忆的重要器官



他的记忆，和金鱼一样，只要“七秒”就会消失。他的生命是一段又一段的空白，没有过去，没有未来。

实现可解释、更鲁棒和更通用的人工智能

数据利用、知识引导与能力学习



数据驱动的机器学习方法

先验与知识引导

从浅层计算到深度神经推理

从单纯依赖于数据驱动模型到数据驱动与知识引导相结合

从领域任务驱动智能到更为通用条件下的强人工智能（从经验中学习）



谢谢大家