|  |  |  |  |
| --- | --- | --- | --- |
| 模型 | 挑战 | 问题 | 文章 |
| Neural Turing Machine | 1. RNN是图灵完备的，理论上可以模拟任何过程（算法），但是实际中是很难完成的。 | 1. 拓展了RNN，能够解决一些算法任务。 2. In contrast to most models of working memory, our architecture can learn to use its working memory instead of deploying a fixed set of procedures over symbolic data.(NTM可以学习到如何使用记忆—读写操作，而不是固定几类操作—栈，队列，等等) 3. It’s design to solve tasks that require the application of approximate rules to “rapidly-created variables 4. 提供寻址机制，用于存储和调用相对简单的原子的数据 5. Variable-binding and Variable-length | Neural Turing Machine |
| Neural GPUs | 1. NTMs are not parallel and are hard to train due to their large depth when unfolded. 2. In its basic form, the entire input is encoded into a single fixed-size vector, so the model cannot generalize to inputs much longer than this fixed capacity. 3. In the best case one would desire a neural network model able to learn arbitrarily complex algorithms given enough resources. 4. NTM use of soft attention requires accessing the entire memory in order to simulate 1 step of computation, which introduces substantial over head | 1. Highly parallel which makes it easier to train and efficient to run(解决了左边提出的各种问题，主要是关于NTM的) 2. This(学习算法) opens the way to use neural networks in domains that were previously only addressed by discrete methods, such as program synthesis | Neural GPUs Learn Algorithms |
| RL-NTM | 1. Many important existing interfaces, such as databases and search engines, are discrete.(除了memory，还有许多外部接口可以使用，比如搜索引擎等，主要是提出不可微的外部接口-离散的外部接口) 2. 离散的接口不能直接使用反向传播训练模型 3. 该模型不是非常powerful,因为很难训练而且只能解决相对简单的问题。解决的任务中，没有超线性 | 1. Discrete Interfaces. Investigate the following discrete interfaces: a memory Type, an input Type, and an output Type 2. 结合reinforce algorithm解决离散外部接口的问题 3. 不同的看待memory的观点---Interface | Reinforcement Learning Neural Turing Machines-Revised |
| Stack RNN | 1. Artifacial Intelligence 2. 当前模型的局限，举例说明了XOR问题和深度学习流行的原因 3. We find that these regularities are difficult to learn even for some advanced deep learning motheds, such as recurrent networks. 4. Wiles and Elman show that simple recurrent networks are able to learn sequences of the form a[n]b[n] and generalize on a limited range of n.（没有学到模式，只是记住了训练中出现的数据） 5. 对于这些简单的算法问题，NTM是否太过于复杂 6. 直观上，连续的模型可能不如离散的模型表现好，未来会两种模型展开研究。 7. 更灵活的记忆模块，比如循环，随机取值等 8. 复杂的算法可以通过简单的算法组合而成，是否可以设计一个模型来完成这一过程 9. 与该文章相同的模型改造方法也可以应用于其他模型，比如multi-dimensional tape | 1. Inreease the learning capabilities of recurrent nets by allowing them to learn how to control an infinite structured memory. 2. 比十九世纪的一些工作的结果来的好，模型比现在的简单。 3. 提出了stack和list两种外部记忆模型，解决算法学习问题。 | Inferring Algorithmic Patterns with Stack-Augmented Recurrent Nets |
| Grid LSTM | 1. Each layer cannot dynamically select or ignore its inputs, it seems attractive to generalize the advantages of LSTM to deep computation. | 1. The network provides a unified way of using LSTM for both deep and sequential computation | Grid Long Short-term memory |
| Memory Networks  (总结) | 1. Most machine learning models lack an easy way to read and write to part of a long-term memory component, and to combine this seamlessly with inference 2. 视觉和听觉任务中，都需要long term memory, 比如看电影并且回答电影中的问题。 | 1. 给出记忆网路的一个框架 2. Efficient memory via hashing 3. QA任务 | Memory Networks |
| DNC | 1. ANN are limited in their ability to represent variables and data structures and to store data over long timescales without interference 2. As the memory demands of a task increase, these networks cannot allocate new storage dynamically, nor easily learn algorithms that act independently of the values realized by task variables 3. NTM has no mechanism to ensure that blocks of allocated memory do not overlap and interfere 4. NTM has no way of freeing locations 5. …… | 1. DNC有能力解决一些复杂的问题，而没有外部记忆模块的神经网络无法解决这些问题 2. 解决NTM的几个问题 |  |
| Pointer Networks | 1. 似乎没有做什么贡献 | 1. The size of the output dictionary depends on the length of the input sequence | Pointer Networks |
| DMN |  | 1. 主要贡献就是提出一种模型，应用于自然语言处理的任务，并取得不错的效果 | Ask me anything: Dynamic Memory Network for Natural Language Processing |
|  |  |  | Neural Random-Access Machines |
|  |  |  | Learning Efficient Algorithms with Attentive Memory |
| DNGPU | 1. To learn algorithms for tasks with unknown solution. 2. Neural GPUs 只在一部分的模型上有泛化能力 3. 如何泛化到任意长度，为什么有泛化能力仍是待解决的问题。 | 1. 对Neural GPU提出了一些改进，减少了训练时间，增加了泛化能力 2. Introduce a technique of general applicability to use hard nonlinearities with saturation cost 3. Introduce a technique of diagonal gates that can be applied to active-memory models 4. 第一个可以学习十进制乘法的结构 | Improving the Neural GPU Architecture for Algorithm Learning |
|  |  |  | Extensions and Limitations of The Neural GPU |
|  |  |  | Lstm recurrent networks learn simple context-free and context-sensitive |
|  |  |  | Dimensions in Program Synthesis |
|  |  |  | Learning to execute |