Summary of papers related to Memory and Multi-tasking

Ganga Meghanath : IIT Madras

Abstract

This is a one page compilation of brief overviews of some papers relevant to multi-tasking in Reinforcement learning setup as well as agents requiring memory. Additional papers are referenced.

1. FRMQN in Minecraft

(Oh et al., 2016) introduces an architecture (adapted from (Sukhbaatar et al., 2015))that performs well in partially observable environments with delayed rewards using high dimensional input observations. The final DQN-based architecture has a separate short-term and long-term memory inter-connected by feedback loops. The architecture generalizes well over temporally extended frames and is able to retrieve memory blocks through time while taking appropriate actions for the current time step.

2. Neural Map

(Parisotto & Salakhutdinov, 2017) introduces an architecture similar to (Oh et al., 2016), but uses an adaptive and sparse write operation to memory (reducing frequent overwriting) selective to the agent's current location, storing information over long time periods in 2D and 3D environments.

3. Pathnet

(Fernando et al., 2017) uses a single large neural network capable of multi-tasking, continual learning and transfers by evolving pathways (evaluated using fitness function) corresponding to each task that could have shared paths from other tasks. The learning of tasks is done in a sequential fashion whereas pathway for a single task can be learned sequentially or in parallel using an A3C (Mnih et al., 2016).

4. Overcoming catastrophic forgetting

(Kirkpatrick et al., 2017) In order to reduce catastrophic forgetting in a multi-task learning setup, the weight updates relevant to previous tasks are selectively constrained using *elastic weight consolidation* to allow continual learning.

5. Active Sampling

(Sharma et al., 2017) A single A3C(Mnih et al., 2016) agent is trained *online* on mutliple tasks in a sequential, occasionally periodic fashion by sampling the difficult tasks relatively more frequently using 3 different techniques.

6. Actor-Mimic

(Teh et al., 2017)AMN uses Policy distillation technique(Rusu et al., 2015) to inherit task-specific policies from expert DQN(Mnih et al., 2015) networks into a single DQN network to perform multiple tasks and enhance transfer (feature regression from experts).

7. Distral

(Teh et al., 2017)Distill and transfer learning is a joint multi-task learning and transfer setup where individual task-specific workers learn in a constrained fashion by a shared policy- distilled to be the centroid of task specific policies.

8. DiGrad

(Dewangan et al., 2018) Differential Policy Gradient derived from DDPG(Lillicrap et al., 2015), is used for joint multi-task learning in the context of robotic systems having continuous action space with shared sets of actions across tasks using an A3C based framework (Mnih et al., 2016).

9. Memory-based Control with RNN

(Heess et al., 2015a)A short term memory for recent sensory inputs coupled with a long term memory is incorporated into a recurrent framework to enable continuous control in partially observable environemnts, by an algorithm derived from DPG(Silver et al., 2014) and SVG(Heess et al., 2015b).

10. Memory in Autonomous Exploration

(Dooraki & Lee, 2018) For better autonomous exploration and obstacle avoidance, memory has been incorporated into the agent : *Memory-based Multilayer Q-Network*; learning from scratch using a model-free off-policy method and maintains a linear sort-term as well as a long-term memory.

11. Episodic Memory

(Gershman & Daw, 2017)The paper suggests that memory is an integral part of a learning system, enabling data-efficient learning in sparse conditions and long term dependencies between actions and rewards; and stresses on the importance of an episodic memory(relating to events) in an RL agent.

12. Joint Many-Task Model

(Hashimoto et al., 2016)The model was proposed in the context of NLP; the depth of the model is increased with task complexity with short cut connections from higher level to lower level to enable task hierarchy, giving outputs at different layers of the model.

Note : Other relevant works include Attend, Adapt and Transfer(Rajendran et al., 2015), Continuous Memory States(Zhang et al., 2016), Progressive Learning(Berseth et al., 2018), Universal Value Function Approximators(Schaul et al., 2015), Deep Relationship Networks(Long & Wang, 2015), Sluice Networks(Ruder et al., 2017), etc.

References

Berseth, Glen, Xie, Cheng, Cernek, Paul, and Van de Panne, Michiel. Progressive reinforcement learning with distillation for multi-skilled motion control. *arXiv preprint arXiv:1802.04765*, 2018.

Calandriello, Daniele, Lazaric, Alessandro, and Restelli, Marcello. Sparse multi-task reinforcement learning. In *Advances in Neural Information Processing Systems*, pp. 819–827, 2014.

Dewangan, Parijat, Phaniteja, S, Krishna, K Madhava, Sarkar, Abhishek, and Ravindran, Balaraman. Digrad: Multi-task reinforcement learning with shared actions. *arXiv preprint arXiv:1802.10463*, 2018.

Dooraki, Amir Ramezani and Lee, Deok Jin. Memory-based reinforcement learning algorithm for autonomous exploration in unknown environment. *International Journal of Advanced Robotic Systems*, 15(3):1729881418775849, 2018.

Fernando, Chrisantha, Banarse, Dylan, Blundell, Charles, Zwols, Yori, Ha, David, Rusu, Andrei A, Pritzel, Alexander, and Wierstra, Daan. Pathnet: Evolution channels gradient descent in super neural networks. *arXiv preprint arXiv:1701.08734*, 2017.

Gershman, Samuel J and Daw, Nathaniel D. Reinforcement learning and episodic memory in humans and animals: an integrative framework. *Annual review of psychology*, 68:101–128, 2017.

Hashimoto, Kazuma, Xiong, Caiming, Tsuruoka, Yoshimasa, and Socher, Richard. A joint many-task model: Growing a neural network for multiple nlp tasks. *arXiv preprint arXiv:1611.01587*, 2016.

Heess, Nicolas, Hunt, Jonathan J, Lillicrap, Timothy P, and Silver, David. Memory-based control with recurrent neural networks. *arXiv preprint arXiv:1512.04455*, 2015a.

Heess, Nicolas, Wayne, Gregory, Silver, David, Lillicrap, Tim, Erez, Tom, and Tassa, Yuval. Learning continuous control policies by stochastic value gradients. In *Advances in Neural Information Processing Systems*, pp. 2944–2952, 2015b.

Kirkpatrick, James, Pascanu, Razvan, Rabinowitz, Neil, Veness, Joel, Desjardins, Guillaume, Rusu, Andrei A, Milan, Kieran, Quan, John, Ramalho, Tiago, Grabska-Barwinska, Agnieszka, et al. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences*, pp. 201611835, 2017.

Lillicrap, Timothy P, Hunt, Jonathan J, Pritzel, Alexander, Heess, Nicolas, Erez, Tom, Tassa, Yuval, Silver, David, and Wierstra, Daan. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

Long, Mingsheng and Wang, Jianmin. Learning multiple tasks with deep relationship networks. *CoRR, abs/1506.02117*, 3, 2015.

Mnih, Volodymyr, Kavukcuoglu, Koray, Silver, David, Rusu, Andrei A, Veness, Joel, Bellemare, Marc G, Graves, Alex, Riedmiller, Martin, Fidjeland, Andreas K, Ostrovski, Georg, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529, 2015.

Mnih, Volodymyr, Badia, Adria Puigdomenech, Mirza, Mehdi, Graves, Alex, Lillicrap, Timothy, Harley, Tim, Silver, David, and Kavukcuoglu, Koray. Asynchronous methods for deep reinforcement learning. In *International conference on machine learning*, pp. 1928–1937, 2016.

Oh, Junhyuk, Chockalingam, Valliappa, Singh, Satinder, and Lee, Honglak. Control of memory, active perception, and action in minecraft. *arXiv preprint arXiv:1605.09128*, 2016.

Parisotto, Emilio and Salakhutdinov, Ruslan. Neural map: Structured memory for deep reinforcement learning. *arXiv preprint arXiv:1702.08360*, 2017.

Rajendran, Janarthanan, Lakshminarayanan, Aravind S, Khapra, Mitesh M, Prasanna, P, and Ravindran, Balaraman. Attend, adapt and transfer: Attentive deep architecture for adaptive transfer from multiple sources in the same domain. *arXiv preprint arXiv:1510.02879*, 2015.

Ruder, Sebastian, Bingel, Joachim, Augenstein, Isabelle, and Søgaard, Anders. Learning what to share between loosely related tasks. *arXiv preprint arXiv:1705.08142*, 2017.

Rusu, Andrei A, Colmenarejo, Sergio Gomez, Gulcehre, Caglar, Desjardins, Guillaume, Kirkpatrick, James, Pascanu, Razvan, Mnih, Volodymyr, Kavukcuoglu, Koray, and Hadsell, Raia. Policy distillation. *arXiv preprint arXiv:1511.06295*, 2015.

Schaul, Tom, Horgan, Daniel, Gregor, Karol, and Silver, David. Universal value function approximators. In *International Conference on Machine Learning*, pp. 1312–1320, 2015.

Sharma, Sahil, Jha, Ashutosh, Hegde, Parikshit, and Ravindran, Balaraman. Learning to multi-task by active sampling. *arXiv preprint arXiv:1702.06053*, 2017.

Silver, David, Lever, Guy, Heess, Nicolas, Degris, Thomas, Wierstra, Daan, and Riedmiller, Martin. Deterministic policy gradient algorithms. In *ICML*, 2014.

Sukhbaatar, Sainbayar, Weston, Jason, Fergus, Rob, et al. End-to-end memory networks. In *Advances in neural information processing systems*, pp. 2440–2448, 2015.

Teh, Yee, Bapst, Victor, Czarnecki, Wojciech M, Quan, John, Kirkpatrick, James, Hadsell, Raia, Heess, Nicolas, and Pascanu, Razvan. Distral: Robust multitask reinforcement learning. In *Advances in Neural Information Processing Systems*, pp. 4496–4506, 2017.

Zhang, Marvin, McCarthy, Zoe, Finn, Chelsea, Levine, Sergey, and Abbeel, Pieter. Learning deep neural network policies with continuous memory states. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 520–527. IEEE, 2016.