# Literature overview

Reinforcement Learning

***Intention of the document:***

The intention of the document is to provide an overview of potential sources, located through reading, and at the same time keeping track of the material that has been read. The references marked with ‘*(read)’* are references that has been read carefully, while those without are potential sources.

**Application area: *Neuroscience***

* Hawkins and Kandal, 1984
* Gelperin, Hopfield, and Tank, 1985
* Tesauro, 1986
* Byrne, Gingrich, and Baxter, 1990
* Friston et al., 1994

Some of the earliest work

* **Marblestone, A. H. et al. (2016): Toward an integration of deep learning and neuroscience.**
* **Song, H.F. et al. (2017): Reward-based training of recurrent neural networks for cognitive and value-based tasks**.
* **Yasmins and DiCarlo (2016): Using goal-driven deep learning models to understand sensory cortex.**
* **Daw et al. (2005): Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioural control. Computation and Systems. Nature Neuroscience.** (read).
* **Daw, N. D. (2012). Model-based reinforcement learning as cognitive search: Neurocomputational theories.** (read)

**Application area: *General RL***

* **Barto, Sutton, and Anderson 1983**
* **Sutton, 1984**
* **Anderson, 1986**

Some of the first work on actor-critic.

* **Sutton, 1988**

Introduction of TD-lambda.

* **Ian Witten, 1977**
* **Ian Witten, 1976a**

Earliest work on TD.

* **Tesauro, G. (1995): Temporal Difference Learning and Backgammon.** (read)

The perhaps best-known success story of RL.

* **Tsitsiklis and Van Roy (1997): An Analysis of Temporal-Difference Learning with Function Approximation.**

Uses linear function approximations.

* **Bellemare, M. G. et al. (2013): The Arcade Learning Environment: An Evaluation Platform for General Agents. (read)**

The introduction of Atari games as a benchmark platform for RL agents.

* **Sutton, R. S. and Barto A. G. (2018). Reinforcement Learning – An introduction. Second Edition, MIT Press, Cambridge, MA.**

The main source

Chapters read so far:

1 and 15

* **Daw, N. D. and Frank, M. J. (2009): Reinforcement learning and higher level cognition: Introduction to special issue. Elsevier.** (read)
* **Juliani, A. (2016): Bridging Cognitive Science and Reinforcement Learning Part 1: Enactivism. Online:** [**https://medium.com/beyond-intelligence/bridging-cognitive-science-and-reinforcement-learning-part-1-enactivism-601af34ef122**](https://medium.com/beyond-intelligence/bridging-cognitive-science-and-reinforcement-learning-part-1-enactivism-601af34ef122) **Accessed: 30/05/2019.** (read)
* **Botvinick et al. (2019). Reinforcement Learning, Fast and Slow. Trends in Cognitive Sciences, Vol. 23, Issue 5, P408-422, May 01.** (read)
* **Yang, Y. (2019). Reinforcement Learning: Value Function Approach. Slides from *Multi-Agent Artificial Intelligence,* a *taught* course at UCL. (read).**
* **Zhang, H. (2019). Reinforcement Learning: Policy-based Methods. Slides from *Multi-Agent Artificial Intelligence,* a *taught* course at UCL. (read).**
* **Silver, D. (2016). RL Course by David Silver. Online:** <https://www.youtube.com/playlist?list=PLzuuYNsE1EZAXYR4FJ75jcJseBmo4KQ9-> Accessed 07/02/2019.
* **Long-Ji Lin. (1993). Reinforcement learning for robots using neural networks. Technical report, DTIC Document.**

**Application area: *Deep RL***

* Mnih V. et al (2013): Playing Atari with deep reinforcement learning. (read)

Some of the first work using deep structures.

* **Rodriguez, J. (2018): Decentralised and Scalable Multi-Agent Reinforcement Learning.** (read)

The future.

* **Arleo et al. (2004): Cognitive navigation based on nonuniform Gabor space sampling, unsupervised growing networks, and reinforcement learning. IEEE Transactions on Neural Networks, Vol. 15, No. 3.** (read)
* **Juliani, A. et al. (2018). Unity: A General Platform for Intelligent Agents.** (read)
* Schulman et al., (2017). Proximal Policy Optimization. Online: <https://openai.com/blog/openai-baselines-ppo/> Accessed: 31/05/2019.
* Mnih V, Kavukcuoglu K, Silver D, et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540): 529-533.
* Silver, David, et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature 529.7587: 484-489.

**Application area: Navigation**

* **Faust, A. and Francis A. (2019).** [**Long-Range Robotic Navigation via Automated Reinforcement Learning**](http://ai.googleblog.com/2019/02/long-range-robotic-navigation-via.html)**. Google AI Blog.**

**Made up by three papers:**

* **Faust, A. et al. (2018). PRM-RL: Long-range Robotic Navigation Tasks by Combining Reinforcement Learning and Sampling-based Planning. *IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, Australia (2018), pp. 5113-5120.**
* **Francis, A. et al. (2019). Long-Range Indoor Navigation with PRM-RL.**
* **Chiang, H-T. L. et al. (2019). Learning Navigation Behaviors End-to-End With AutoRL.** [**IEEE Robotics and Automation Letters**](https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=7083369)**( Volume: 4 ,**[**Issue: 2**](https://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=8581687)**, April 2019 ).**

**Comment:**

State of the art – beautiful!

* **Kavraki, L. E. et al. (1996). Probabilistic roadmaps for path planning in high-dimensional configuration spaces. IEEE Transactions on Robotics and Automation (Volume: 12, Issue: 4, Aug 1996).**

**Comment:**

Keyword: Static workspace

* **Zhang, J. et al. (2017). Deep Reinforcement Learning with Successor Features for Navigation across Similar Environments.** (Read).

**Comment:**

Generalisable framework, both in terms of changing goals and environments). However, the real improvement appears (read more carefully before postulating that) to only be present after the initial training, which still justifies trying to improve the initial training (Potentially through spatial cues, as proposed).

They assume similarity in the presence of the spatial objects, which is not necessarily true for different cities.

* **Yen, G. G. and Hickey, T. W. (2004). Reinforcement learning algorithms for robotic navigation in dynamic environments.** [**ISA Transactions**](https://www.sciencedirect.com/science/journal/00190578)**,** [**Volume 43, Issue 2**](https://www.sciencedirect.com/science/journal/00190578/43/2)**, April 2004, Pages 217-230.** (read)

**Comments:**

Essentially what I intended to do.

* **Zuo, B. et al. (2014). A reinforcement learning based robotic navigation system.** [**2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)**](https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6960119)**.** (read).

**Comments:**

They employ Q-learning embedded in a small robot, which using (internal) sensors keeps track of the environment around it and navigate according to the received information.

**Application area: Uncertainty combined with reinforcement learning**

* **Kahn et al. (2017). Uncertainty-Aware Reinforcement Learning for Collision Avoidance.**

**Comments:**

They propose an extension to the traditional RL problem, by included a collision prediction

model, to determine the speed of movement for the agents. The task of the collision

prediction model is to regulate the speed when passing through novel environments, to

avoid potentially catastrophic crashes doing training.

~~This does not seem relevant for the intended purpose, because the idea is to model cities as~~

~~accurate as possible in the training environment.~~

Could be used in the low-level part of the hierarchical implementation.

* **Lütjens et al. (2018). Safe Reinforcement Learning with Model Uncertainty Estimates.**

Their work is intended for what you could call micro navigation, implying dealing with unseen behaviour of potentially colliding objects. Using their own words:

*“Note that our work targets model uncertainty estimates because they potentially reveal sections of the test data where training data was sparse and a model could fail to generalize …”* [Introduction].

* **Garcia and Fernández (2015). A Comprehensive Survey on Safe Reinforcement Learning.** **Journal of Machine Learning Research 16 (2015). 1437-1480.**

They survey different ways of incorporating risk in the standard reinforcement learning problem, and they present three papers which are especially interesting:

**The three papers:**

Driessens and Dzeroski (2004)

Martín H. and Lope (2009)

Song et al. (2012)

**Start points for literature review:**

RL dating back to 1970, with some of the initial break throughs originating there.

Application areas spanning a wide stretching range of topics, ranging from finance to neuroscience and over to playing games.

**Structure of literature review:**

* Coming to where we are today
* Deep RL
* Focus area
* Software
  + Previously state-of-the-art benchmarking
  + The future and the applied (Unity)

**Random thoughts:**

* Could one reuse layers (weights) from existing D-NN (used for function approximation) when training a new agent, to increase generalisation of RL in general?