Obtaining Approximately Admissible Heuristic Functions through Deep Reinforcement Learning and A* Search

Forest Agostinelli, Stephen McAleer, Alexander Shmakov, Roy Fox, Marco Valtorta, Biplav Srivastava, Pierre Baldi

AI INSTITUTE #AIISC UNIVERSITY OF SOUTH CAROLINA

1

Abstract

- Deep reinforcement learning methods have been used to learn heuristic functions for A* search.
- These heuristic functions are not guaranteed to be admissible.
- We develop a domain agnostic method that corrects an inadmissible heuristic called approximately admissible conversion.
- On the 15-puzzle and 24-puzzle, our method produces a heuristic function that is empirically admissible in over 99.99% of cases and finds a shortest path for 100% of test cases.

2

Approximately Admissible Conversion

Approximately admissible conversion uses the fact that incomplete runs of A* can be used to estimate a lower bound on the cost of a shortest path (see paper for proof). The method is as follows:

1. A subset of the state space is obtained and an estimate of a lower bound on the cost of a shortest path is initialized to zero for each state
2. The inadmissible heuristic function is adjusted so that it does not overestimate the lower bound of any state in the subset
3. A* search is then performed. A solution does not need to be found as the costs of expanded nodes are used to update the lower bound.
4. This procedure then goes back to step 2 until the lower bounds stop increasing.

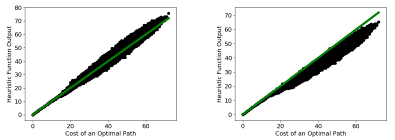
3

Experiments

- Using approximate value iteration, we train a deep neural network to be a heuristic function for the 15-puzzle and 24-puzzle.
- We find the cost of a shortest path for 1 million states from the 15-puzzle and compare it to the output of the adjusted heuristic function.
- We solve 500 test states for both the 15-puzzle and 24-puzzle and compare to the cost of a shortest path.

4

Admissibility Results



(a) Before approximately admissible conversion

(b) After approximately admissible conversion

X-axis: cost of a shortest path. Y-axis: Output of heuristic function. All points above the green line are inadmissible

- Before: 71.37% inadmissible
- After: 0.0019% inadmissible
- Heuristic is still highly correlated with the cost of a shortest path and thus informative.

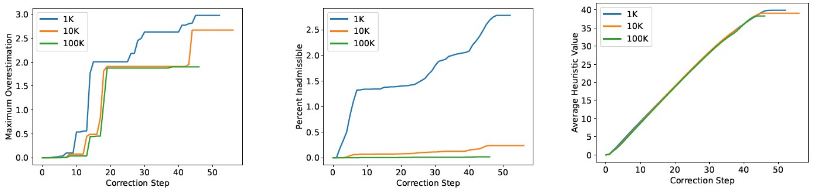
5

Shortest Path Results

- We use a batched version of A* search to take advantage of parallelism provided by GPUs.
- With approximate admissible conversion, a shortest path is found 100% of the time for both the 15-puzzle and 24-puzzle.
- Without it, the percentage of the time a shortest path is found can decrease by up to 31.5%.

6

Effect of Subset Size



(a) Max overestimation (b) Percent inadmissible (c) Average heuristic value

- A larger subset decreases inadmissibility while maintaining an informative heuristic.

7

Discussion

- Deep neural networks have the ability to learn informative heuristic functions in a domain-independent fashion.
- This can greatly increase the ease with which search methods such as A* search can be applied.
- Many real-world problems in fields such, as chemical synthesis, robotics, and quantum computing have connections to deep reinforcement learning and can benefit from shortest path solutions to increase efficiency and reduce resource consumption.

8

References

- Agostinelli, F.; McAleer, S.; Shmakov, A.; and Baldi, P. 2019. Solving the Rubik's cube with deep reinforcement learning and search. Nature Machine Intelligence 1(8): 356–363.
- Bertsekas, D. P.; and Tsitsiklis, J. N. 1996. Neuro-dynamic programming. Athena Scientific. ISBN 1-886529-10-8.
- Harris, L. R. 1974. The heuristic search under conditions of error. Artificial Intelligence 5(3): 217–234.
- Hart, P. E.; Nilsson, N. J.; and Raphael, B. 1968. A formal basis for the heuristic determination of minimum cost paths. IEEE transactions on Systems Science and Cybernetics 4(2): 100–107.
- Zhang, Y.-H.; Zheng, P.-L.; Zhang, Y.; and Deng, D.-L. 2020. Topological Quantum Compiling with Reinforcement Learning. Physical Review Letters 125(17): 170501.