

# Solving the rubiks cube with inverse RL

People who try to solve a rubiks cube without prior knowledge about the algorithms to do so, mostly find it way harder than they imagined. That is because it is rather unintuitive, often you have to break up matching patterns in order to advance forward to solving the whole cube. There are also different ways to get there.

If we translate this into to RL realm this means that rewards are sparse, having a huge state space with only one goal state.[1]

The algorithm (=policy) to solve the cube however, is well known. It does not lead to an optimal solution but guarantees to find one. Therefore I want to propose to use inverse reinforcement learning[2] on this problem to learn the reward function from the policy and maybe even try to use the learned reward function to improve the policy.

There are gyms on easy (2x2)[3] and standard (3x3)[4] cubes.

## References

- [1] F. Agostinelli, S. McAleer, A. Shmakov, and P. Baldi, "Solving the Rubik's cube with deep reinforcement learning and search," *Nat. Mach. Intell.*, vol. 1, no. 8, pp. 356–363, Aug. 2019, doi: 10.1038/s42256-019-0070-z.
- [2] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proceedings of the twenty-first international conference on Machine learning*, New York, NY, USA, Jul. 2004, p. 1. doi: 10.1145/1015330.1015430.
- [3] K. Hukmani, *RubiksCubeGym*. 2022. Accessed: Jul. 14, 2022. [Online]. Available: <https://github.com/DoubleGremlin181/RubiksCubeGym>
- [4] R. Chiu, *gym-Rubiks-Cube*. 2022. Accessed: Jul. 14, 2022. [Online]. Available: <https://github.com/RobinChiu/gym-Rubiks-Cube>

Feedback [rgratz@uos.de](mailto:rgratz@uos.de)