

Soft Actor Critic

Oliver, Leon, Thilo Röthemeyer

13. April 2021

Contents

Part 1

Part 2

Conclusion

- Performance im Vergleich zu TD3 und A2C
- Vor- und Nachteile von SAC

Performance

- SAC soll stabiler laufen als TD3 Algorithmen
- Testumgebung gym

Ziel der Experimente

- Stabilität und Sample Komplexität im Vergleich zu anderen Algorithmen
 - kontinuierliche Aufgaben
- OpenAI gym und rllab
 - schwierig für off-policy Algorithmen
 - Parametertuning bei komplexeren Aufgaben schwierig

Vergleich zu anderen Algorithmen

- SAC
 - mean action
- DDPG
- PPO
- SQL mit zwei Q Funktionen
 - evaluated with exploration noise
- TD3

Ergebnisse

- f

Zusammenfassung

- soft actor critic vorgestellt
 - off policy
 - Entropiemaximierung verbessert Stabilität
 - Besser als state of the art Algorithmen