Thesis Progress Form
CHARLES DARWIN UNIVERSITY
College of Engineering, IT, and Environment

Name: Shane Reynolds

Unit: ENG720

Title: Automatic generation control of a two area power system using deep reinforcement learning

Supervisors: Charles Yeo & Stefanija Klaric

Time & Date: August 24, 2020 @ 1pm

1   Progress since last meeting

- Completed experimental trial using experimental replay - agent performance did not beat PID control

- Examined idea presented in Yan and Xu regarding using supervised learning in conjunction with a PID controller to pre-train the agent. This would be time consuming to implement, so experimented with an alternative approach. Took the best trained agent from previous exeriemnts (i.e. closest to PID level control that has been achieved) and loaded the weights into a new DDPG training session, allowing the the agent to train for a further 5000 episodes. This did not result in a successfully trained agent.

- Analysed the currently specified reward function to double check that additive elements of the function were not acting against each other (i.e. the agent is punished for using the generators, but also punished for allowing frequency deviations). Re-ran original experiments using a modified reward function such that minimising the frequency deviations was given a higher priority than low generator use. The result was no meaningful improvement on frequency control outcomes; however, a worse control scheme was observed. Concluded that original reward function specification was okay.

2 Discussion Points

- Outlined recent progress to CY, as described above.

- Outlined new experiments to CY, as detailed below.

3 Plan until the next meeting

- Another variation on the Yan and Xu paper would be to fill the agent replay buffer with 50% of data from a PID controller at all times. That is allow the agent to use its own policy for 50% of the time and use a PID policy for 50% of the time. This would flood the agent experience replay memory with expert controller examples, and help the agent to learn a suitable policy. The idea is that the agent would take the most prominent aspects of the

- Revisit code base to ensure that DDPG and training algorithms have been implemented correctly. Look at using an off the shelf implementation sourced from Github for DDPG and check results against previous experiments.

- Revisit implementation of two area power system environment and re-confirm that this has been implemented correctly.

_____

Supervisor

August 24, 2020