A Reinforcement Learning Approach for Blood Glucose Control without Carbohydrate Estimation

Mirjam Oppliger

Supervisors: Prof. Dr. Stavroula Mougiakakou, PD Dr. Lorenzo Brigato
Institution: University of Bern, ARTORG Center for Biomedical Engineering Research
Examiners: Prof. Dr. Stavroula Mougiakakou, PD Dr. Lorenzo Brigato

Introduction

Diabetes mellitus Type I is a chronic metabolic disease characterized by high levels of blood glucose (BG) due to the body's inability to produce insulin. Affected people require exogenous insulin administration. Advancements in technology reduced the burden of the disease by introducing insulin pumps and continuous glucose measurement (CGM) sensors. These systems build the basis for an artificial pancreas (AP) with the goal to fully automate insulin infusion without user interaction. While modern systems already automatically adapt insulin dosage according to the measured CGM value, all available systems require user initiated insulin boluses when carbohydrates (CHO) are consumed. This thesis focuses on the implementation and evaluation of reinforcement learning (RL) agents eliminating the need for CHO estimation.

Materials and Methods

To train RL agents a Type I Diabetes Simulator with ten in silico patients was used. Different RL agents exists, however, for this project the proximal policy optimization (PPO) algorithm was chosen for its sample efficiency, real-time decision-making, and convergence properties. Furthermore, policy optimization allows for optimization of complex policies. The basic implementation was adapted and extended to meet the complex task. Different modifications were tested in multiple configurations. Applied modifications include changes to the network architectures of the actor and critic network, feature selection, and feature extraction.

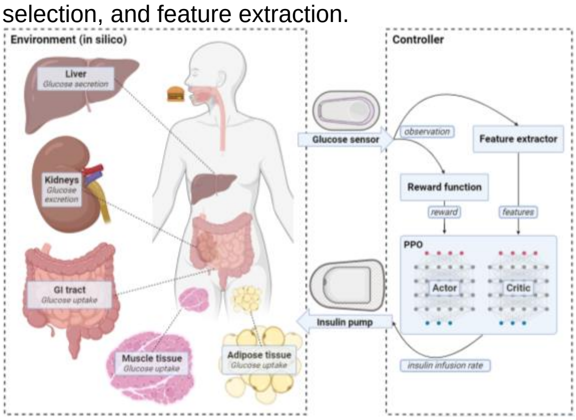

Fig. 1 Setup of the AP. The environment implements the glucose-insulin system which interacts with the sensor, controller and insulin pump. The controller computes the insulin dosage based on the reward from the environment and the CGM measurements of the sensor. Created with BioRender.com

The PPO agents without having CHO information as input were compared to basal-bolus (BB) controllers with different meal misestimations and to a PPO agent having access to CHO information contained in meals consumed. Outcome measures included time spend in different BG ranges, risk for severe hypoglycemia, control-grid variability analysis, and qualitative analysis of BG curves.

Results

The most successful agent trained achieved satisfactory results (73.3% time in range TIR) with a tendency to overshoot insulin dosage and thereby sending the patient into hypoglycemia. Agents having access to CHO information generally achieve a higher TIR (86.6% BB and 82.7% PPO with CHO). Agents generally struggle to grasp the prolongedness of the insulin effects accurately.

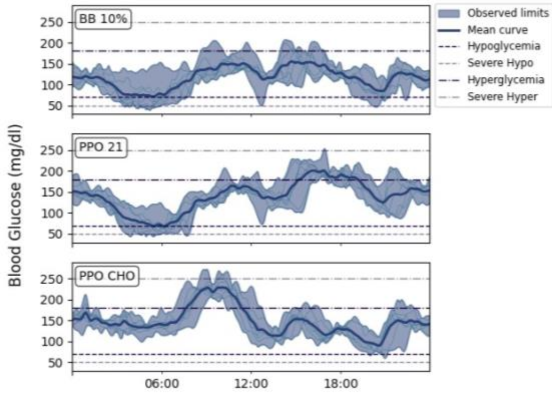

Fig. 2 Daily average profiles for different controllers.

Discussion

While these results show promising advances toward a fully automated AP, further research and validation is necessary. Agents proposed in this thesis do not consider the criticality of the correct insulin dosage and lead to dangerous BG values leading to severe hypoglycemia in many of the in-silico patients.

References

M. Tauschmann and R. Hovorka, "Technology in the management of type 1 diabetes mellitus - current status and future prospects," Nature Reviews. Endocrinology, vol. 14, no. 8, pp. 464–475, Aug. 2018

Acknowledgements

I extend my sincere thanks to Dr. Lorenzo Brigato and Prof. Dr Stavroula Mougiakakou for their mentorship, providing me with valuable feedback and guidance during this thesis.