TBMI26 – Computer Assignment Reports  
Reinforcement Learning

Deadline – Mars 12 2018

Author/-s:

In order to pass the assignment you will need to answer the following questions and upload the document to LISAM. You will also need to upload all code in .m-file format. If you meet the deadline we correct the report within one week after the deadline. Otherwise we give no guarantees when we have time.

1. **Define the V- and Q-function given an arbitrary policy as well as a given optimal policy (See lectures/classes).**
2. **Define a learning rule for the Q-function (Theory, see lectures/classes).**
3. **Describe your implementation, especially how you hinder the robot from exiting through the borders of a world.**
4. **Describe the differences between the worlds explored by the robot. Any surprises?**
5. **For each world: Plot the V-function, i.e. how do you get to the goal from each position.**
6. **For each world: describe the key observations you have made with respect to parameter choices. Provide documentation of the parameters you have used for each figure! A good rule is to provide each figure with a caption. Plot policies and the V-function for appropriate worlds to the extent you find appropriate in order to explain what you have done and learned during the assignment.**
7. **What would happen if we where to only use Dijkstra's shortest path finding algorithm in the ''Suddenly Irritating blob'' world? What about in the static ''Irritating blob'' world?**
8. **Include an in-depth description of the to/from HG worlds (world 3 and 4). What happens on the way from HG? How and why can this problem be solved with Q-learning? Which path does the robot prefer, and why?**
9. **Can you think of any application where reinforcement learning could be of practical use? A hint is to use the Internet.**
10. **How does the different parameters () influence learning and appearance of the Q- and V-functions?**