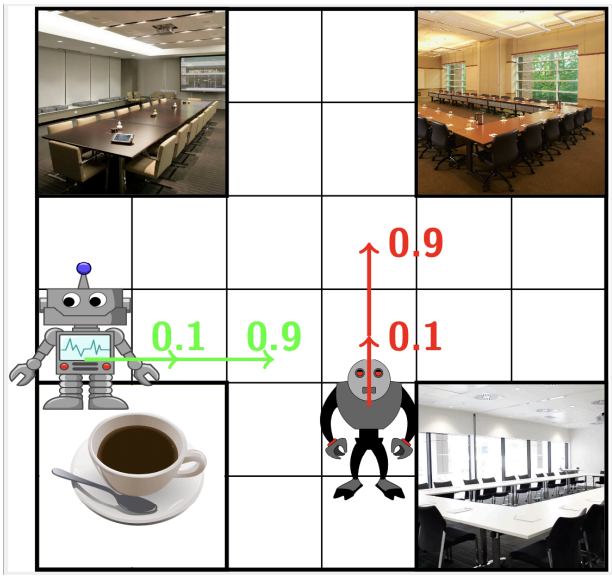
Algorithmic collusion for omega-regular properties in grid-world scenarios

Supervisor: Janusz Meylahn (j.m.meylahn@utwente.nl),
Ernst Moritz Hahn (e.m.hahn@utwente.nl)

Consider a scenario of multiple agents, each of which is following their own interest and is using reinforcement learning (RL) to maximize their profit. Under certain conditions, collusion may occur, that is, the RL algorithms might implicitly negotiate contracts with other agents.

In this proposed thesis, you are going to research collusion in scenarios where agents operate on a grid-world and follow an omega-regular objective. Omega-regular properties are properties which can be described using omega-regular automata or logics. They can describe more complex behaviours than reward signals. For instance, consider the example below:



The first robot tries to distribute coffee from the coffee room to the other three rooms. This behaviour can be described as **Pmax** [**G** (**F** coffee $\wedge$ **XF** room1 $\wedge$ **XF** room2 $\wedge$ **XF** room3)]. The second robot wants to move between two rooms, expressible as **Pmax** [**G** (**F** room1 $\wedge$ **F** room2)]. To efficiently perform these tasks, profit from negotiating with each other, so as to avoid each other's path. It would be convenient for each robot to ignore the other robot completely rather than moving out of the way, because in this way it might be able to distribute coffee more quickly. However, in this case, the other robot would reciprocally do the same, if according RL algorithms are used.

There is an active area of using RL for handling omega-regular properties. The novel aspect of this thesis is the combination with collusion. The challenges of this thesis are

- implementing appropriate transformations from omega-regular properties to rewards,
- implementing appropriate 2-player RL algorithms on top of the above,
- evaluating the results, and
- providing an outlook for a theoretical foundation of these results.