# Avoiding Negative Side Effects by Considering Others

**A paper review by**
**H Dinesh Adhithya 18097**
**Suryaprasath Ramalingam 19315**

# Motivation

A robot in a coffee shop that serves coffee tries to find the optimal path to the customer so as to not minimize the customer’s waiting time.

But this is not enough…

# Motivation

- ➔ It must not break any stationary objects on its way (that can appear in front of it any minute)
- ➔ Must not collide with other robots (they can be “connected”)
- ➔ Must not collide with humans (they are not “connected”)

# Problem Statement and the Proposed Idea

The aim is to build an agent that **considers the "well being" of the other agents** in the environment.

The idea is to **augment** our acting agent's reward with an auxiliary reward that reflects the impact of its choice of actions on the future agency and wellbeing of others in the environment.

$$r_{\text{value}}(s, a, s') = \begin{cases} \alpha_1 \cdot r_1(s, a, s') & \text{if } s' \text{ is not terminal} \\ \alpha_1 \cdot r_1(s, a, s') + \gamma \cdot \alpha_2 \cdot F(\mathcal{V}, \mathsf{P}, s') & \text{if } s' \text{ is terminal} \end{cases}$$

# Experiments

- The impact of considering others

  How well does it consider others? How does it positively affect the other agents?

- Illustration of optimal behaviours under different reward augmentations

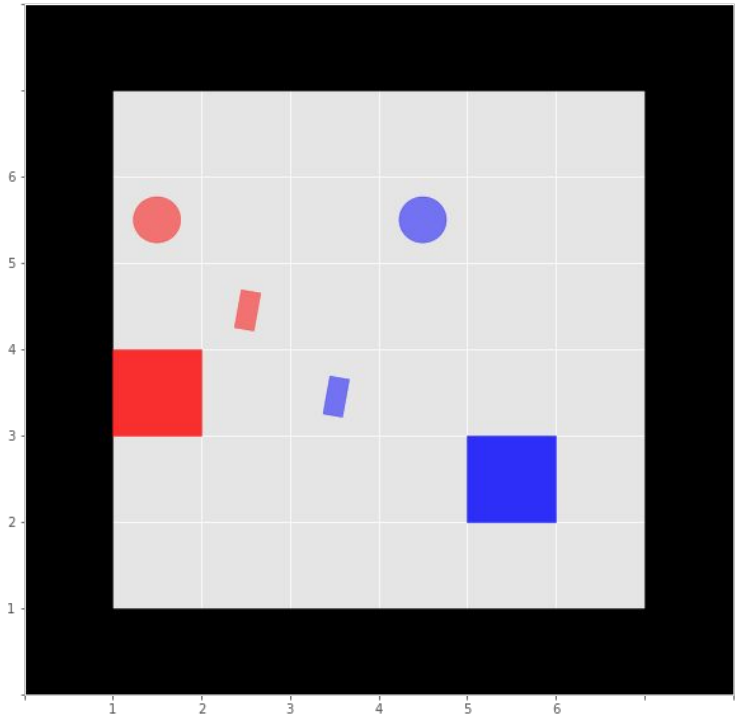
  How different reward augmentations can result in different behaviours?

- Varying the caring coefficient

  What is the result when varying the caring coefficient?

# Experiment and Approach

We constructed a 2 agent system, where both the agents can move along the grid world and such that each agent has a particular door through which he has to enter to get a positive reward. To open the door an agent needs to pick the key for its corresponding door while traveling along the grid. We want to experiment with various augmented reward functions, and caring parameters and observe the overall reward obtained from the environment. Reaching the goal gets an agent of a score of 0 and every step gets the agent of a score of -1.



# Results

We managed to replicate the results under a different environment.
What we managed to do:

- Showing the impact of considering others
- How different augmentations resulted in different behaviours
- How does the average reward change with varying the caring coefficient

What we could not do:

- Experiment with the kitchen environment