Politecnico di Torino
PIC4SeR
King's College London

Learning Social Heuristics for Human-Aware Path Planning

Andrea Eirale
Matteo Leonetti
Marcello Chiaberge

PRECISION AGRICULTURE

SMART CITY SEARCH & RESCUE

SERVICE ROBOTICS FOR WELLBEING

CULTURAL HERITAGE

UNDERWATER

SPACE

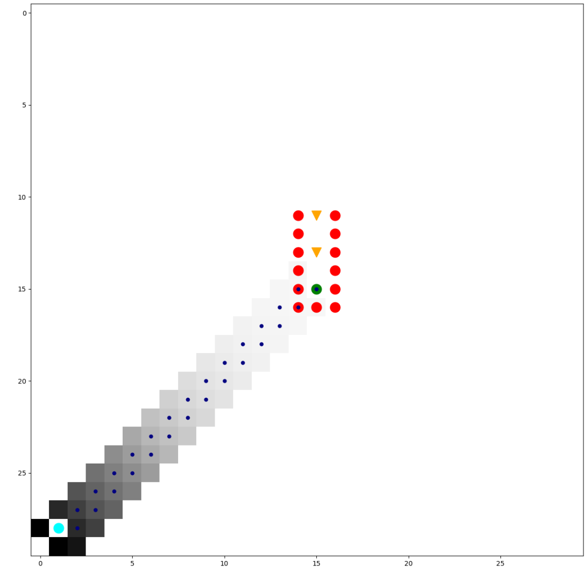
Motivation

In the context of social navigation, the robot should not only respect others' **personal spaces**, but also comply with specific **social norms**

Classic path planning algorithms focus on navigation cost and heuristic (subscripted with n) and do not take **social aspects** into account

$$f(s) = g_n(s) + h_n(s)$$

A series of **socially accepted behaviours** can be learnt and integrated with **classical path planning** at no additional planning costs



Methodology adopted

We propose **Heuristic Planning with Learned Social Value** (**HPLSV**), a method to learn a **value function** encapsulating the cost of **social navigation** and use it as an additional heuristic in heuristic-search path planning

The objective function is modified to consider **social aspects ad contexts**

$$f(s) = g_n(s) + h_n(s) + w(g_s(s) + h_s(s))$$

Where the subscript n represent the **navigation component** of the objective and with the subscript s the **social component** of the objective

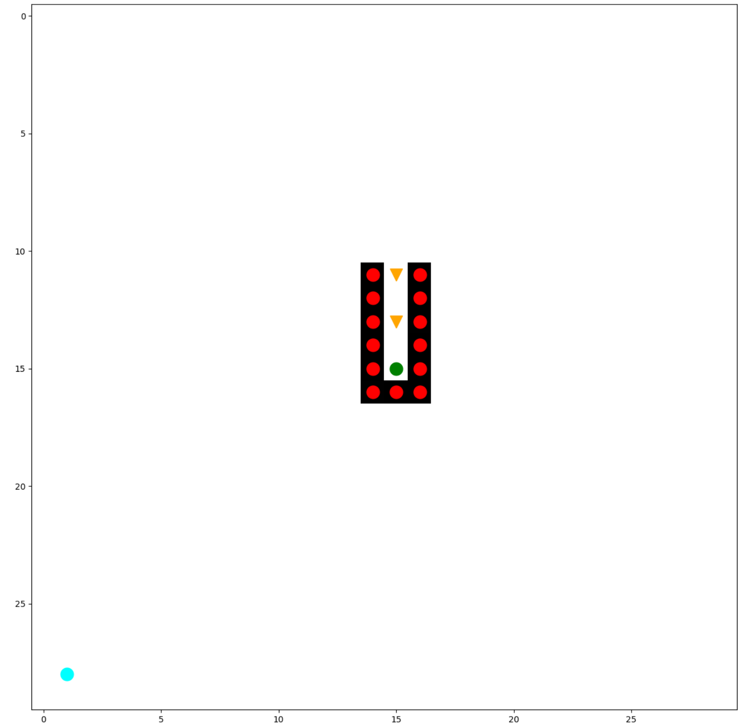
As a proof of concept, for this work we consider the problem of **following a queue**

Methodology adopted

We train an **RL agent** to maximize a reward function $r_T = r_n + r_s$, and learn a value function Q_T

At the same time it learn a **second value function** Q_s, which does not affect the behavioural policy

The value function Q_T is discarded and the heuristic h_s is extracted from the social value function Q_s



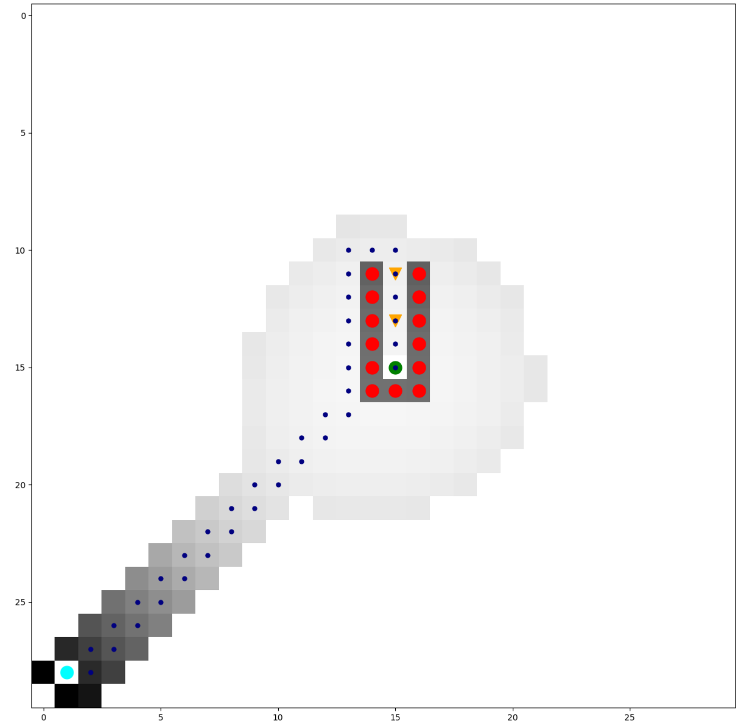
Experimentations and results

At first, the agent is trained in a **demo gridmap environment**

We define **virtual obstacles** as a continuous box surrounding people and goal. If the agent hits the virtual obstacle, it is equivalent to **cutting the queue**

To test and validate the model, it is integrated with a classic $A*$ planner

The planner is now able to recognize the social scenario and extract an **optimal path** to the goal which passes through the end of the queue, avoiding to cut it

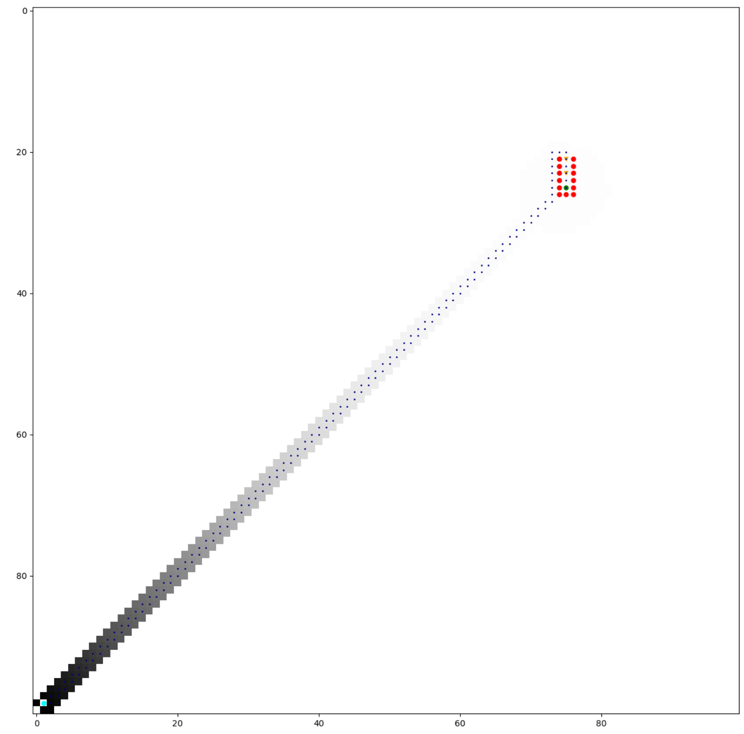


Experimentations and results

The **ego-centric representation** of the heuristic makes it independent of coordinates and allows the planner to work in a much larger testing environment

The **social component** of the objective does not activate (the expected social cost of actions is 0) until the agent gets near the goal

This allows the traditional planner to compute an **optimal trajectory** towards the goal without any other interference until the **social scenario** is reached

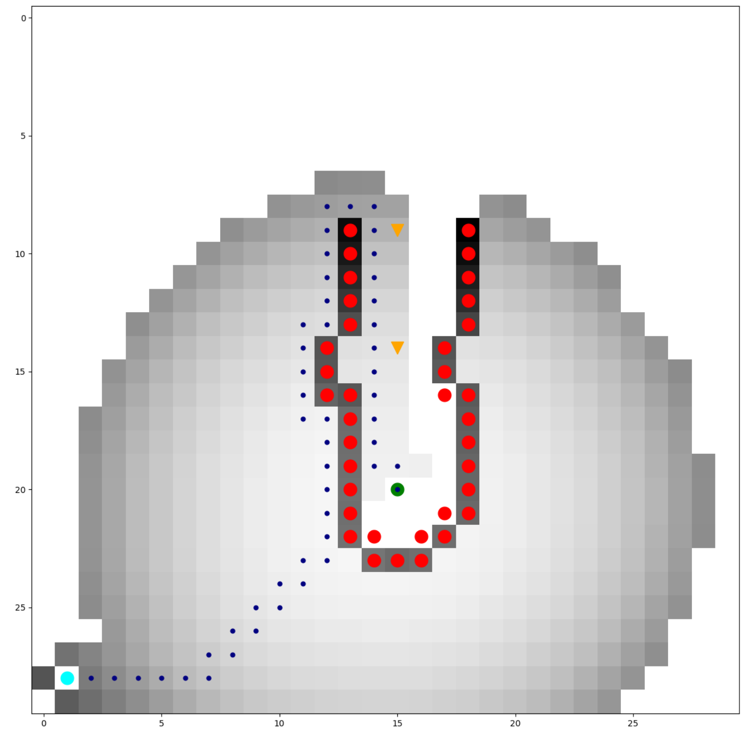


Experimentations and results

A series of different **continuous environments** are obtained from a Gazebo simulation, to train and test the agent in more **realistic conditions**

The new obtained continuous environments are then **discretized** and used to retrain the previous DRL agent

Similarly to before, the model is integrated with the A* planner. The agent is now able to **generalize** on various environments never seen before



Conclusions and Future works

For these preliminary results, we developed a proof of concept on queue following, but intend to **extend the methodology** to more, if not all, human activities

In this work we assumed to know the **exact position and orientation** of all and only the people in the queue. We do not expect the heuristic function to be able to solve this problem end-to-end, directly for robot sensory inputs

Another strong assumption is the a priori knowledge of the **social cost function**, represented in training with virtual obstacles. In general, it would be interesting to naturally learn it from **real human-robot interactions**.