# Team Edinferno Open Challenge Description: Learning to score penalties against reactive goalkeepers

Aris Valtazanos, Subramanian Ramoorthy, Chris Towell, Efstathios Vafeias
School of Informatics, University of Edinburgh

June 4, 2012

The existing limitations in sensing and robot detection make it difficult for SPL robots to reason and interact with their adversaries. Currently, these constraints limit robots to treating other players as stochastically moving obstacles, with considerable uncertainty in their perceived state. However, as these problems are gradually addressed and overcome, it becomes possible to collect richer information on the states and actions of adversarial agents. More importantly, it becomes to possible to make progress towards using formal tools from the autonomous decision making literature (e.g. Partially Observable Markov Decision Processes - POMDPs), which can be used to learn a model of the adversary's behaviour in a principled manner.

In this year's Open Challenge, we will present an application of a POMDP-based algorithm to the penalty shootout problem against a reactive goalkeeper. Our aim is to show how a robot can form a model of its opponent's responses to its own actions, and learn to form policies that are likely to satisfy its own strategic goal.

The demonstration will feature two robots: one striker and one goalkeeper. The striker is only allowed straight – non-directional – kicks, but he can adjust his orientation to shoot towards the edges of the goal. The goalkeeper is not allowed to dive to save the ball, and can only side-step to a better blocking position along the goal line. Thus, the striker seeks to deceive the goalkeeper through appropriate changes in his orientation, into moving to a different side of the goal than the one he is going to kick to.

To overcome the uncertainty in pose estimation, both robots are provided with their own initial position and orientation, and that of their adversary. Then, at each time frame, the robots exchange their latest odometry-based sensed movements, and use them to update their positions. This joint information provides both players with a noisy estimate of their adversary's state, which can be then used in their decision making.

Due to the tight time constraints of the Open Challenge, we will not demonstrate a full learning procedure. Instead, the striker will use policies that have been learned offline to play against two different types of adversaries: a reactive goalkeeper who follows a fixed policy to block the ball, and a joystick-based goalkeeper who is controlled remotely by a human operator. Thus, we seek to demonstrate how the striker can flexibly learn to adapt to unknown adversaries with different characteristics.