Learning to Play Sequential Games versus Unknown Opponents
Pier Giuseppe Sessa, Ilija Bogunovic, Maryam Kamgarpour, Andreas Krause

ICML Workshop on Real World Experiment Design and Active Learning, 18th July 2020
Summary of our work

• Repeated **Sequential Game** setup:
  At each round $t$, 
  • Learner picks action: $x_t$
  • Opponent responds with: $y_t = b(x_t, \theta_t)$
  • Learner obtains reward: $r(x_t, y_t)$

• Learner’s regret: 
  $$R(T) := \max_{x \in \mathcal{X}} \sum_{t=1}^{T} r(x, b(x, \theta_t)) - \sum_{t=1}^{T} r(x_t, y_t)$$

• Novel algorithm **StackelUCB**, with *sublinear* regret guarantees for the Learner.

• **Key Idea**: Sequentially learn the response function $b(\cdot, \cdot)$ via **kernel ridge regression** + employ online learning strategies by using an optimistic estimate
Experiments

- **Routing vehicles in Traffic Networks:**

  ![Time averaged Regret](chart1.png)

  - StackelUCB leads to low regret and reduces the congestion in the network.

- **Wildlife Protection against Poaching Activities:**

  ![Obtained Rewards](chart2.png)

  - StackelUCB discovers the optimal patrol strategy online after ~60 iterations.

- **Check out our paper for more details!**