# SHORT TERM SCIENTIFIC MISSION (STSM) SCIENTIFIC REPORT

**This report is submitted for approval by the STSM applicant to the STSM coordinator**

**Action number:** CA16228
**STSM title:** Overtaking optimality in stochastic games with classical computer science objective functions
**STSM start and end date:** 25/12/2018 to 08/01/2019
**Grantee name:** János Flesch

## PURPOSE OF THE STSM:

We consider zero-sum stochastic games of perfect information in which the payoff function is given by a reachability condition (or by a safety condition). Even though in these games optimal strategies are known to exist, certain optimal strategies clearly outperform other ones on long but finite horizons. This is calling for a refinement of optimality in these games.

Our approach is to introduce an overtaking criterion to the payoff function, which allows us to compare optimal strategies. Those optimal strategies that are the best with respect to this overtaking criterion are called overtaking optimal. Some of the main goals are to analyze when these stochastic games admit such overtakingly optimal strategies, how to find such a strategy, and how simple these overtaking optimal strategies can be. An interesting an important special case is when there is only one player in the game, i.e. the special case of Markov decision processes.

## DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS

First, we considered Markov decision processes (MDP) with a reachability condition. In such an MDP, the player's goal is to reach a certain state. Even though it is known that the player has optimal strategies, certain optimal strategies clearly outperform other ones on long but finite horizons, as examples demonstrate. To address this issue, we defined a refinement of optimality for such MDPs. We called this refinement overtaking optimality. As it turned out, this refinement leads to some interesting results and examples. We can define overtaking optimality in a similar way for MDP with a safety objective. In such an MDP the player's goal is to never visit a certain state.

As a next step, we extended the definition of overtaking optimality to zero-sum stochastic games of perfect information where player 1's goal is to reach a certain state and player 2's goal is to prevent it. Here the refinement leads to a pair of optimal strategies such that each player's strategy is overtaking optimal in the MDP that is obtained by fixing the opponent's strategy.

The main question we examined in detail was when such zero-sum stochastic games (or MDPs) admit such pairs of optimal strategies. In addition, when these can be chosen to be simple, e.g. pure or stationary.

## DESCRIPTION OF THE MAIN RESULTS OBTAINED

We obtained the following results:

**1.** We say that a zero-sum stochastic game of perfect information has quasi-deterministic transitions if, for each state and action, there is at most one state other than the target state to which transition can occur with positive probability. We showed that every quasi-deterministic zero-sum stochastic game of perfect information admits a pair of pure stationary strategies such that, given his opponent's strategy, neither player can change his strategy individually to an overtakingly better strategy. Our proof method is also providing an algorithm to find such a pair of strategies.

**2**. The result in item 1 is sharp in the following sense:
**2.1.** We constructed a quasi-deterministic MDP with reachability condition in which the player has no overtaking optimal strategy at all (all pure strategies are beaten by all mixed strategies, and each mixed strategy is beaten by another mixed strategy).
**2.2.** We also constructed a (non-quasi-deterministic) MDP with reachability condition in which for every pure strategy there is an overtakingly better pure strategy.
We remark that these two MDPs are constructed in a very specific way, and so they are not "generic".

**3**. In the MDPs mentioned in item 2, if the player can even use finitely additive strategies, and not only the usual countably additive ones, then it turns out that the player does have an overtaking optimal strategy. We also know that this existence result should hold in larger generality. (We still need to check some details to see under which conditions this will hold exactly.)

**FUTURE COLLABORATIONS (if applicable)**

The main question remaining is the following: is it true that an overtaking optimal strategy always exists in generic MDPs or in generic zero-sum games of perfect information, as described above? We already know that this may require a careful examination of the speed of convergence of the probabilities that the target state is reached in less than t steps, under certain strategies. We plan to investigate this problem further.

In fact, it would also be interesting to consider classical payoff functions in computer science that are more complicated than reachability or safety conditions. One question is how to define overtaking optimality for such payoff functions and whether or when Markov decision problems or zero-sum games of perfect information admit such optimal strategies. We have some initial ideas and we may investigate this in a follow-up project.