**Research project offer**

**Location**: ISAE SUPAERO, Toulouse, France  
**Department**: Department of Complex Systems Engineering (DISC)  
**Research group**: ADO  
**Supervisor**: Emmanuel Rachelson, Paul-Antoine Le Tolguenec  
**Email**: emmanuel.rachelson@isae-supaero.fr, paul-antoine.le-tolguenec@isae-supaero.fr

### OFFER DESCRIPTION

**Title**: Stabilizing exploration with uncertainty  
**Proposed duration and period**: May 1 to Nov 1, 2023

**Context**  
In deep reinforcement learning, exploration is essential for the development of robust policies. However, when using a parametric neural network to represent policies, maintaining continuous exploration can be challenging due to the inherent instabilities emerging from the training of these networks. In this internship, we will study various regularization mechanisms aimed at stabilizing training methods for exploration in deep reinforcement learning.

**Possibility to continue with a PhD (Yes/No)**: Yes

### REQUIRED APPLICANT PROFILE AND SKILLS

<table>
<thead>
<tr>
<th>Study Level</th>
<th>Undergraduate students (3rd or 4th year)</th>
<th>Master students (2nd year)</th>
<th>PhD students</th>
</tr>
</thead>
</table>

**Objectives and work**  
Recent studies propose uncertainty measures that enable efficient exploration in generic state spaces, such as RND [1] and NGU [2]. A challenge of these methods lies in maintaining a precise equilibrium between efficient exploration and exploitation, which can be difficult to achieve or sustain. In our study, we will investigate two regularization mechanisms which could stabilize these methods. One mechanism will be based on successor state estimation [3], while the other will draw from the field of convex optimization.


**Required profile and skills**  
Experience with deep learning, strong mathematical skills, and the Python programming language are all highly recommended.