

# Planejamento Estocástico e Aprendizado por Reforço

**Código do projeto:** VF01

**Responsável:** Prof. Valdinei Freire da Silva

**Linha de pesquisa:** Inteligência de Sistemas

**Número de vagas:** 2

## Descrição geral

A área de Planejamento Estocástico considera o problema de decisões sequenciais em ambientes probabilísticos, enquanto Aprendizado por Reforço considera o problema no qual um agente interage com um ambiente desconhecido no qual recompensas e penalidades são emitidas pelo ambiente de acordo com as decisões do agente no mesmo. O agente deve aprender como obter recompensas e evitar penalidades por meio de estratégias baseadas em tentativa e erro, que em geral é muito lento.

Esse projeto tem como objetivos: (i) contribuir com algoritmos para critérios de decisão com equidade e tornar políticas interpretáveis para seres humanos, de modo que as preferências com relação ao risco possam ser acessadas pelos mesmos e auxiliem em uma decisão informada, e (ii) explorar formalismos que permitam acelerar o aprendizado de novas tarefas utilizando conhecimento informado ou adquirido em tarefas anteriores.

## Perfil desejado

Conhecimento de formalismos (probabilidade, álgebra e cálculo) e lógica de programação.

## Referências

Bradley Hayes and Julie Shah. Improving Robot Controller Transparency Through Autonomous Policy Explanation, 2017.

Dan Amir and Ofra Amir. HIGHLIGHTS: Summarizing Agent Behavior to People, 2018.

Leila Amgoud, Henri Prade. Using arguments for making and explaining decisions, 2009.

Valdinei Freire, Karina Valdivia Delgado, Willy Reis. An Exact Algorithm to make a Trade-off between Cost and Probability in SSPs, 2019.

Isabella Kuo, Valdinei Freire. Probability-to-Goal and Expected Cost Trade-Off in Stochastic Shortest Path, 2021.

Rediet Abebe, Kira Goldner. Mechanism design for social good, 2018. Rediet Abebe, Jon Kleinberg, S Matthew Weinberg. Subsidy allocations in the presence of income shocks, 2020.

Okan Dukkanci, Özlem Karsub, Bahar Y. Kara. Planning sustainable routes: Economic, environmental and welfare concerns, 2021.

Xuejing Zheng, Chao Yu, Chen Chen, Jianye Hao, Hankz Hankui Zhuo. Lifelong Reinforcement Learning with Temporal Logic Formulas and Reward Machines, 2021.

Roy Fox, Sanjay Krishnan, Ion Stoica, Ken Goldberg. Multi-Level Discovery of Deep Options, 2017.

Pashootan Vaezipoor, Andrew C. Li, Rodrigo Toro Icarte, Sheila McIlraith. LTL2Action Generalizing LTL Instructions for Multi-Task RL, 2021.