

Agentes BDI e Aprendizagem: um mapeamento sistemático e utilização com a biblioteca MASP Y

Felipe Merenda Izidorio, Alexandre L. L. Mellado,
André Pinz Borges, Gleifer Vaz Alves

Universidade Tecnológica Federal do Paraná - Campus Ponta Grossa

16 de agosto de 2024



Sumário

- 1 Introdução
- 2 Mapeamento Sistemático da Literatura
- 3 MASPYP: Aprendizagem e Sistema BDI
- 4 MASPYP: Exemplo de Utilização
- 5 Considerações Finais

Introdução: Conceitos

- Agentes Inteligentes:
 - Tomada de decisão em sistemas autônomos;
- Aprendizagem por Reforço (AR):
 - Baseados no Modelo de Decisão de Markov e nas Equações de Bellman;
 - Algoritmos de AR aprimoram decisões através de interações repetidas com o ambiente;
 - Solução de desafios complexos: Veículos Autônomos (VAs);

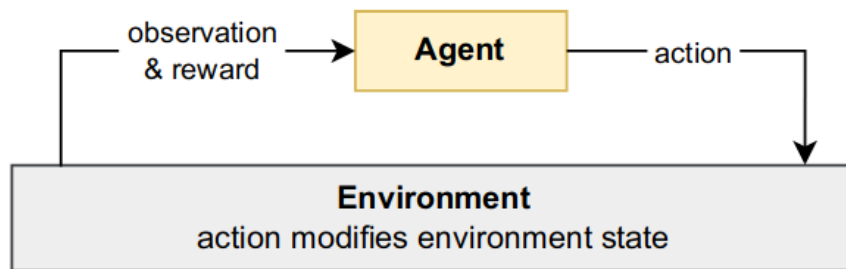


Figura 1: Loop de aprendizado por Reforço.

Introdução: Objetivos

- Iniciativas do grupo de pesquisa da UTFPR, LaCA-IS, visam soluções baseadas em agentes inteligentes para problemas envolvendo trânsito, estacionamento inteligente, veículos autônomos, etc;
- Realizar um mapeamento sistemático na área de aprendizado por reforço, com foco em veículos autônomos;
- Descrever uma extensão da biblioteca MASPY;
- Identificar algoritmos mais utilizados e métodos comuns de validação.

Mapeamento Sistemático da Literatura

- Metodologia de Kitchenham e Chartes;
- Parsifal: <https://parsif.al/>
- Busca por algoritmos de aprendizagem na condução de VAs e encontrar possíveis lacunas dentro da área;

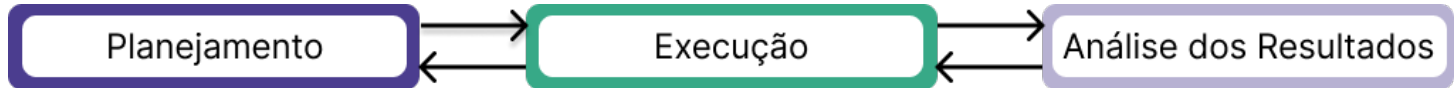


Figura 2: Processo de Mapeamento

Planejamento do Mapeamento

- Questões de pesquisa:
 - **QP1.** Quais algoritmos de aprendizagem são usados no trabalho?
 - **QP2.** É realizada uma comparação com outros algoritmos de aprendizagem?
 - **QP3.** Quais métricas de avaliação são utilizadas no trabalho?
 - **QP4.** Qual a técnica predominante nos trabalhos de condução de VAs?
 - **QP5.** Dentro da área de condução de VAs, qual temática é a mais abordada?

- Palavras-chave:

Reinforcement Learning (Decision Making, Deep Reinforcement Learning, Reinforcement Learning Algorithms) e Autonomous Vehicle Navigation (Autonomous Vehicle Driving, Autonomous Vehicle Movement Control, Self-Driving Car Driving, Self-Driving Car Movement Control”)

Planejamento do Mapeamento

- Bases de busca:
 - IEEE Digital Library
 - Science Direct
 - Scopus
- Critérios de inclusão (**CI**) e exclusão (**CE**):
 - **CI1.** Aborda sobre algoritmos de aprendizagem (**CI2.** ...para agentes) (**CI3.** ...para a condução de VAs)
 - **CE1.** Estudos duplicados / **CE2.** Fora de escopo / **CE3.** Não é estudo primário
- A partir desses critérios será possível eleger os trabalhos que serão usados para a etapa de qualidade.

Planejamento do Mapeamento

- Questões de qualidade:
 - **QQ1.** O trabalho aborda sobre algoritmos de aprendizagem por reforço?
 - **QQ2.** O trabalho faz uma comparação com outros algoritmos?
 - **QQ3.** O trabalho se baseia em alguma arquitetura de agentes?
 - **QQ4.** O trabalho aborda sobre condução de VAs?
 - **QQ5.** O trabalho utiliza o(os) algoritmo(os) para conduzir os VAs?
 - **QQ6.** O trabalho faz uma avaliação do(os) algoritmo(os) abordado(os)?
- Respostas possíveis e respectivas pontuações:
 - Sim, **1.0**
 - Parcialmente, **0.5**
 - Não, **0.0**
- A soma dessas pontuações corresponde ao quão útil o trabalho é para o tema abordado.

Execução do Mapeamento

- Após a etapa de Planejamento, foram obtidos **159** trabalhos;
- Aplicando os critérios de inclusão e exclusão, **34** trabalhos foram selecionados para leitura e realização da etapa de avaliação de qualidade;
- Por fim, **22** trabalhos obtiveram uma pontuação de qualidade acima de 3 pontos, servindo de base para a análise dos resultados;

Análise dos Resultados e Conclusões do Mapeamento

- O algoritmo Deep Q-Networks foi o mais utilizado ([Dar et al.], [Tiong et al. 2023], [Josef and Degani 2020], [Cabezas-Olivenza et al. 2023]);
- A comparação com outros algoritmos é crucial para avaliar sua eficácia e serve como guia para estudos futuros;
- Métricas específicas relacionadas ao problema são mais vantajosas para solucioná-lo;
 - Exemplo: **tempo de locomoção antes de uma colisão** ([Liang et al. 2023], [Arvind and Senthilnath 2019], [González-Miranda et al. 2023] e [Taghavifar et al. 2024]);

Análise dos Resultados e Conclusões do Mapeamento

- A utilização de ferramentas e técnicas de outras áreas pode enriquecer o aprendizado dos agentes;
 - **Visão Computacional:** Redes Neurais Convolucionais [Bin Issa et al. 2021]
 - **Psicologia:** *Social Value Orientation* [Qin et al. 2024];
- A área de condução de VAs tem diversas temáticas exploradas nos trabalhos investigados, como:
 - **prevenção de colisão** ([Liang et al. 2023], [Arvind and Senthilnath 2019], [González-Miranda et al. 2023] e [Taghavifar et al. 2024])
 - **controle de estabilidade** ([Liang et al. 2023] e [Cabezas-Olivenza et al. 2023])
 - **velocidade** ([Hartmann et al. 2020])

Análise dos Resultados e Conclusões do Mapeamento

- Dos trabalhos selecionados no mapeamento, nenhum se baseia em alguma Arquitetura de Agentes Inteligentes, evidenciando uma lacuna na área;
- Diante disso, é apresentada a biblioteca MASPYP:
 - uma biblioteca em Python que combina um sistema multiagente BDI com Aprendizado por Reforço;

MASPY: Aprendizagem e Sistema BDI

- Múltiplos agentes podem ser conectados a múltiplas ambientes e estes podem trocar mensagem entre vários canais de comunicação;
- O administrador controla a configuração e execução do sistema, enquanto o aprendizado pode estender as capacidades de agentes.

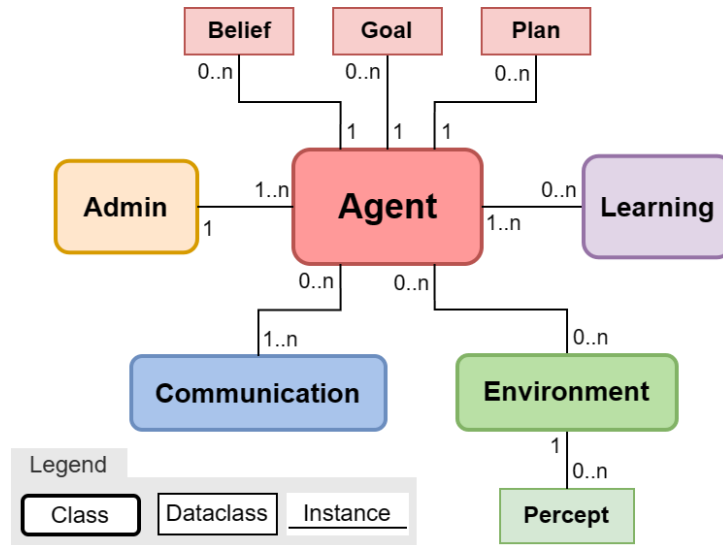


Figura 3: Diagrama de Classes do MASPY.

MASPY: Ciclo de Raciocínio

- Agentes seguem um ciclo de raciocínio BDI, com atualização de crenças e objetivos a partir de percepções e mensagens, para executar planos e ações em ambientes.

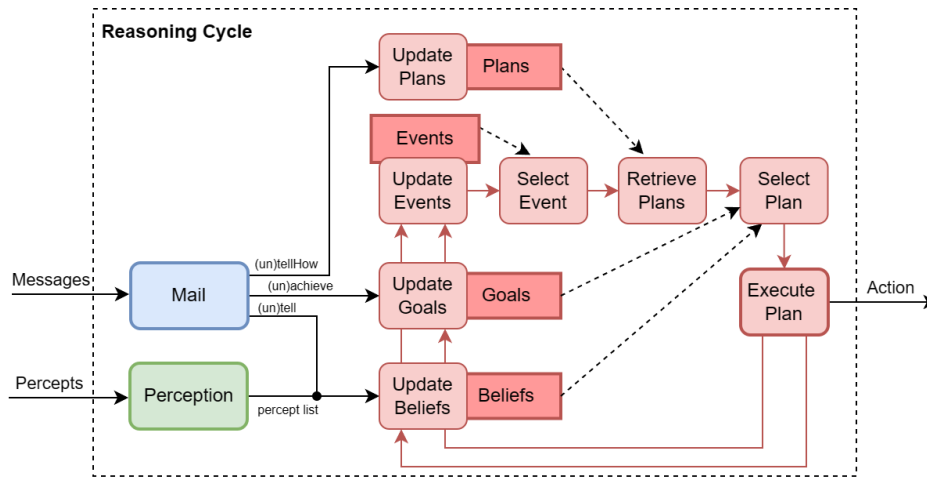


Figura 4: Ciclo de raciocínio BDI dos agentes.

MASPY: Aprendizagem

- Um agente pode utilizar a classe de aprendizagem para expandir suas ações disponíveis;
- Uma instância desta classe contém a modelagem de um ambiente específico e retorna a melhor ação para cada estado.

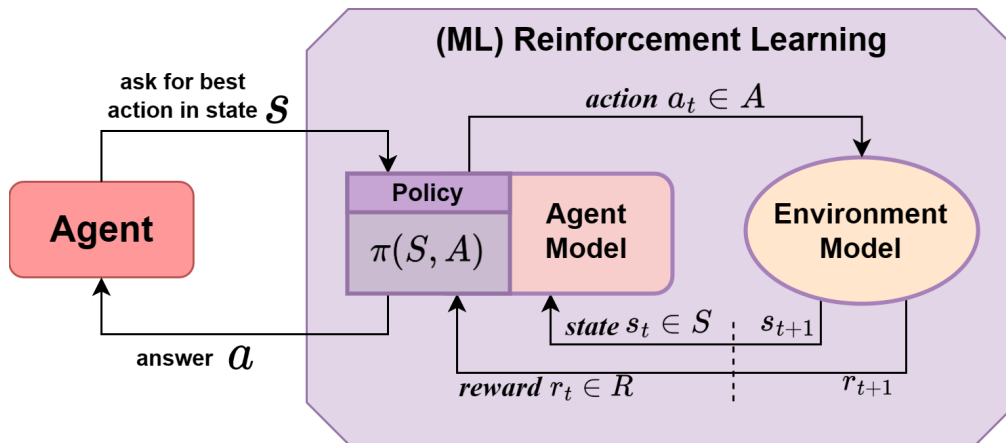
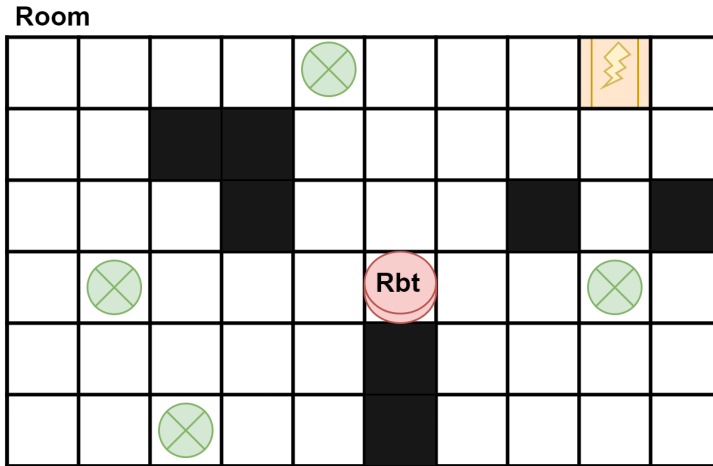


Figura 5: Classe de Aprendizagem e sua utilização por agentes.

MASPY: Exemplo de Utilização



- Agente robô aspirador, ambiente com obstáculos, sujeira e posições alvo.

- A navegação é feita a partir do trajeto aprendido pela classe de aprendizagem;
- A limpeza é feita pelo agente robô aspirador percebendo e interagindo com o ambiente durante seu ciclo de raciocínio.

Considerações Finais

- O mapeamento sistemático identificou elementos relevantes de investigação a respeito de aprendizagem e condução de veículos autônomos;
- Como a necessidade de uma abordagem que faça uso de uma arquitetura específica para programação de agentes (em especial, agentes BDI) com técnicas de aprendizagem.
- Descrição da biblioteca MASPYP com elementos básicos de aprendizagem.

Trabalhos Futuros:

- Validação das capacidades da biblioteca MASPYP;
- Novos algoritmos de aprendizagem;
- Expandir a etapa de aprendizagem no ciclo de raciocínio de agentes.



Bosello, M. (2020). Integrating bdi and reinforcement learning: the case study of autonomous driving.

Integrating BDI and Reinforcement Learning: the Case Study of Autonomous Driving

Supervisor:

Prof. Alessandro Ricci

Co-supervisor:







Prof. Giovanni Pau

Author:

Michael Bosello

Academic Year 2019-2020

Referências

-  Dar, S. A., Palanivel, S., and Geetha, M. K. Autonomous Taxi Driving Environment Using Reinforcement Learning Algorithms. *International Journal of Modern Education and Computer Science*, 14(3):88.
-  Tiong, T., Saad, I., Teo, K. T. K., and Lago, H. B. (2023). Autonomous Vehicle Driving Path Control with Deep Reinforcement Learning. In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 0084–0092, Las Vegas, NV, USA. IEEE.
-  Josef, S. and Degani, A. (2020). Deep Reinforcement Learning for Safe Local Planning of a Ground Vehicle in Unknown Rough Terrain. *IEEE Robotics and Automation Letters*, 5(4):6748–6755. Conference Name: IEEE Robotics and Automation Letters.
-  Cabezas-Olivenza, M., Zulueta, E., Sanchez-Chica, A., Fernandez-Gamiz, U., and TesoFz-Beto no, A. (2023). Stability Analysis for Autonomous Vehicle Navigation Trained over Deep Deterministic Policy Gradient. *Mathematics*, 11(1):132. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.
-  Liang, J., Weerakoon, K., Guan, T., Karapetyan, N., and Manocha, D. (2023). AdaptiveON: Adaptive Outdoor Local Navigation Method for Stable and Reliable Actions. *IEEE Robotics and Automation Letters*, 8(2):648–655.
-  Arvind, C. S. and Senthilnath, J. (2019). Autonomous RL: Autonomous Vehicle Obstacle Avoidance in a Dynamic Environment using MLP-SARSA Reinforcement Learning. In *2019 IEEE 5th International Conference on Mechatronics System and Robots(ICMSR)*, pages 120–124, Singapore. IEEE.

Referências



González-Miranda, O., Miranda, L. A. L., and Ibarra-Zannatha, J. M. (2023). Q-Learning for Autonomous Vehicle Navigation. In *2023 XXV Robotics Mexican Congress (COMRob)*, pages 138–142.



Taghavifar, H., Wei, C., and Taghavifar, L. (2024). Socially Intelligent Reinforcement Learning for Optimal Automated Vehicle Control in Traffic Scenarios. *IEEE Transactions on Automation Science and Engineering*, pages 1–12. Conference Name: IEEE Transactions on Automation Science and Engineering.



Bin Issa, R., Das, M., Rahman, M. S., Barua, M., Rhaman, M. K., Ripon, K. S. N., and Alam, M. G. R. (2021). Double Deep Q-Learning and Faster R-CNN-Based Autonomous Vehicle Navigation and Obstacle Avoidance in Dynamic Environment. *Sensors*, 21(4):1468. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.



Qin, J., Qin, J., Qiu, J., Liu, Q., Li, M., and Ma, Q. (2024). SRL-ORCA: A Socially Aware Multi-Agent Mapless Navigation Algorithm in Complex Dynamic Scenes. *IEEE Robotics and Automation Letters*, 9(1):143–150. Conference Name: IEEE Robotics and Automation Letters.



Hartmann, G., Shiller, Z., and Azaria, A. (2020). Model-Based Reinforcement Learning for Time-Optimal Velocity Control. *IEEE Robotics and Automation Letters*, 5(4):6185–6192. Conference Name: IEEE Robotics and Automation Letters.

Contatos

- Felipe Merenda Izidorio
felipemerenda@alunos.utfpr.edu.br
- Alexandre L. L. Mellado
mellado@alunos.utfpr.edu.br
- André Pinz Borges
apborges@utfpr.edu.br
- Gleifer Vaz Alves
gleifer@utfpr.edu.br
- LaCA - Intelligent Systems
<https://laca-is.github.io/>
<https://github.com/laca-is/MASPY>

