

Aprendizagem por Imitação para Condução Autônoma Segura de Automóveis

Proposta de Dissertação de Mestrado 2020/2021

Orientador

Prof. Eric Aislan Antonelo

Palavras-chave

Aprendizagem por reforço, aprendizagem adversária generativa, inteligência artificial, condução autônoma em ambientes urbanos, minimização de eventos de risco

Introdução

Aprendizagem por imitação é uma abordagem pela qual um modelo caixa-preta (rede neural) é treinado para imitar um especialista usando um conjunto fixo de amostras de pares observação-ação (ou trajetórias) obtidas daquele especialista. A clonagem comportamental (CC) é um tipo de aprendizagem por imitação baseada em um processo de treinamento supervisionado de um modelo usando um grande conjunto de dados. A CC tem sido utilizada para a obtenção de políticas de condução autônoma para veículos, onde as amostras de treinamento são geradas por um conjunto de motoristas humanos. No entanto, o método de CC gera agentes enviesados, que tenderão a possuir certos comportamentos determinados pelo tipo de motorista empregado na geração das amostras. Além disso, a CC sempre sofrerá com erros em cascata (erro que é retroalimentado), pois seu conjunto de treinamento não conterà todas possíveis situações de condução em uma rodovia ou dentro de uma cidade.

Uma nova abordagem chamada *Generative Adversarial Imitation Learning* (GAIL) [1] pode construir políticas de agentes a partir de um conjunto de trajetórias como em CC, mas que são mais robustas do que as obtidas por CC, por permitir que o agente aprenda em interação com o ambiente como em aprendizagem por reforço. Isso faz com que o agente experimente trajetórias inexistentes no conjunto de trajetórias original, o que lhe permite aprender políticas de controle robustas para situações não previstas nos dados iniciais. O GAIL permite:

- obter políticas de agente robustas como em aprendizagem por reforço sem a necessidade da definição de uma função de recompensa;
- ser usado com redes neurais profundas para processar entradas de alta dimensionalidade (imagens capturada por câmera);
- usar um conjunto de treinamento muito menor do que em CC, ou seja, é eficiente em termos do número de amostras (*sample efficient*).

Basicamente, existem dois componentes do GAIL que são treinados iterativamente em um jogo min-max: um classificador discriminativo D é treinado para distinguir entre amostras geradas pela política de aprendizado π e amostras geradas pela política π_E do especialista ou motorista humano (representada pelo conjunto de treinamento rotulado); e a política de aprendizagem π é otimizada para imitar a política π_E do especialista. Matematicamente, GAIL encontra um ponto de sela (π, D) da expressão:

$$\mathbb{E}_{\pi}[\log(D(s, a))] + \mathbb{E}_{\pi_E}[\log(1 - D(s, a))] - \lambda H(\pi) \quad (1)$$

onde $D : S \times A \rightarrow (0, 1)$, S é o espaço de estados, A é o espaço de ações; π_E é a política do especialista; $H(\pi)$ é um regularizador de política controlado por $\lambda \geq 0$. Tanto a política π , quanto o discriminador D podem ser representados por redes neurais profundas.

Para que o agente opere de modo seguro, é necessário adicionar termos sensíveis ao risco à função sendo otimizada [3, 2], de modo que os eventos (ou comportamentos) de baixa qualidade sejam minimizados (chance de acidentes em uma cirurgia ou durante uma viagem de táxi autônomo).

Objetivos

Neste tema, o mestrando atuará na área de aprendizagem por imitação para condução autônoma de veículos, buscando desenvolver:

- experimentos com o simulador CARLA (<http://carla.org/>) para geração de conjuntos de dados (trajetórias do especialista) em ambientes urbanos
- redes neurais profundas treinadas por GAIL para condução autônoma de veículos
- método para treinamento da rede neural minimizando eventos de risco (de fim de cauda da distribuição)

Perfil do candidato

Você tem: habilidade para leitura e escrita na língua inglesa e capacidade de trabalho individual e em grupo; grande motivação para estudos avançados e para aprender a usar simulador de ambientes urbanos CARLA. É desejável ter experiência em frameworks em Python **tensorflow/pytorch** (redes neurais), e base sólida de conhecimento em aprendizagem de máquina (e.g.: treinamento, *overfitting*, validação cruzada, etc.) e aprendizagem por reforço.



Figure 1: Simulador CARLA

Plano de atividades preliminar

O mestrando selecionado se capacitará em temas relevantes como aprendizagem por imitação, redes neurais, simulação de ambientes urbanos e implementará modelos segundo a metodologia proposta. As seguintes atividades estão previstas:

1. **Capacitação em Aprendizagem por Imitação:** Consiste no estudo dos conceitos básicos de aprendizagem de máquina (AM), na revisão da literatura em aprendizagem por imitação, em especial trabalhos envolvendo o método GAIL. Uma implementação deste em Python usando biblioteca existente <https://stable-baselines.readthedocs.io/en/master/modules/gail.html> será realizada para uma tarefa de controle de baixa dimensão denominada cart-pole usando a plataforma OpenAI. (<https://gym.openai.com/envs/CartPole-v0/>). Será comparado com uma implementação de clonagem comportamental.
2. **Revisão da literatura em veículos autônomos:** Engloba o estudo dos trabalhos recentes de navegação autônoma em ambientes urbanos, em particular usando o simulador CARLA (<http://carla.org/>).
3. **Simulação de veículos em ambientes urbanos:** Consiste no estudo, instalação e emprego da plataforma CARLA para simulação de navegação de veículos em ambientes urbanos. Será necessário implementar scripts em Python para coleta de dados de navegação de um motorista no ambiente urbano. Serão definidos o ambiente, os tipos de dados (sensores, atuadores) coletados, bem como a tarefa de navegação (ex.: fazer a volta em um bloco com o automóvel).
4. **Implementação de modelos GAIL para navegação:** Consiste no projeto e implementação de redes neurais que aprendam pelo método GAIL, empregando o conjunto de dados coletados na fase anterior. O treinamento exigirá a implementação de uma conexão com o simulador CARLA em tempo real. A primeira parte corresponde a uma tarefa mais simples usando sensores do tipo *LIDARs*, enquanto a segunda parte empregará imagens de câmera como entrada para o agente. Para cada caso, redes profundas ou convolucionais serão escolhidas e suas configurações definidas.
5. **Implementação do GAIL sensível ao risco para navegação:** No contexto das tarefas e conjunto de dados do passo anterior, consiste em adaptar e implementar o método *Risk-averse Imitation Learning* (RAIL) [3] para minimizar o risco de acidentes comuns em ambientes urbanos e eventos raros de alto custo em geral.
6. **Avaliação e Comparação:** Nesta etapa, os modelos desenvolvidos anteriormente serão avaliados e comparados entre si, considerando métricas de desempenho a serem definidas (duração do percurso, colisão, aceleração média, etc.). A comparação poderá incluir agentes de condução autônoma já disponíveis no simulador CARLA: agente *AutoWare*, e agente *Conditional Imitation Learning*.
7. **Redação de documentos:** Consiste na escrita de artigos científicos relacionados a este tema e aos resultados obtidos e da dissertação de mestrado.

Referências bibliográficas

- [1] Jonathan Ho and Stefano Ermon. Generative adversarial imitation learning. In *Advances in Neural Information Processing Systems*, pages 4565–4573, 2016.
- [2] Jonathan Lacotte, Mohammad Ghavamzadeh, Yinlam Chow, and Marco Pavone. Risk-sensitive generative adversarial imitation learning. In *22nd International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2019.
- [3] Anirban Santara, Abhishek Naik, Balaraman Ravindran, Dipankar Das, Dheevatsa Mudigere, Sasikanth Avancha, and Bharat Kaul. Rail: Risk-averse imitation learning. *arXiv preprint arXiv:1707.06658*, 2017.