


# THÔNG TIN CHUNG CỦA BÁO CÁO

- Link YouTube video của báo cáo (tối đa 5 phút):  
<https://youtu.be/pyoFqoybReU>
- Link slides:  
<https://github.com/dovugiacan14/Traffic-Signal-Control/blob/main/presentation.pdf>
- *Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới*
- *Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in*

<ul style="list-style-type: none"><li>● Họ và Tên: Đỗ Vũ Gia Cần</li><li>● MSSV: 230101073</li></ul> 	<ul style="list-style-type: none"><li>● Lớp: CS2205.CH181</li><li>● Tự đánh giá (điểm tổng kết môn): 8.5/10</li><li>● Số buổi vắng: 1</li><li>● Số câu hỏi QT cá nhân: 3</li><li>● Link Github: <a href="https://github.com/dovugiacan14/Traffic-Signal-Control">https://github.com/dovugiacan14/Traffic-Signal-Control</a></li></ul>
---	---

# ĐỀ CƯƠNG NGHIÊN CỨU

## TÊN ĐỀ TÀI (IN HOA)

XÂY DỰNG CHIẾN LƯỢC ĐIỀU KHIỂN TÍN HIỆU GIAO THÔNG TỰ ĐỘNG  
BẰNG PHƯƠNG PHÁP HỌC TĂNG CƯỜNG

## TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

DEVELOPING AN AUTOMATED TRAFFIC SIGNAL CONTROL USING  
REINFORCEMENT LEARNING

## TÓM TẮT *(Tối đa 400 từ)*

Điều phối giao thông hiệu quả tại các giao lộ được xem là một trong những giải pháp cho việc giảm thiểu tình trạng ùn tắc giao thông tại thành phố lớn, nơi mật độ giao thông ngày càng tăng do sự gia tăng dân số, từ đó bài toán Điều khiển tín hiệu giao thông tự động (TSC) được ra đời với mục tiêu là tìm ra các chiến lược để thay đổi các tín hiệu đèn giao thông hiệu quả theo lưu lượng phương tiện di chuyển qua các giao lộ. Chúng tôi đề xuất áp dụng các thuật toán học tăng cường cho bài toán TSC, đánh giá hiệu quả của các chiến lược được tìm ra bởi các phương pháp này so với các chiến lược truyền thống trước đó thông qua nhiều tình huống giao thông khác nhau được mô phỏng sát với thực tế của bộ mô phỏng giao thông SUMO, dựa trên các tiêu chí đánh giá như: tổng thời gian chờ trung bình của các xe tại các giao lộ, trung bình tổng số xe chờ trên tất cả các làn đường, trung bình tổng độ trễ. Cụ thể, chúng tôi sẽ mô hình hóa bài toán TSC dưới dạng một tác vụ học tăng cường, và dựa trên một bản đồ mô phỏng giao thông qua lại một hoặc nhiều giao lộ cho trước, chúng tôi sẽ cho thuật toán tự động học và tìm ra các chiến lược tối ưu tương ứng với từng tình huống giao thông khác nhau.

## **GIỚI THIỆU** (Tối đa 1 trang A4)

Sự gia tăng dân số không ngừng qua mỗi năm tại các trung tâm kinh tế trọng điểm hay các thành phố, khu đô thị lớn gây ra nhiều tác động xấu đến nhiều mặt đời sống, xã hội, trong đó lĩnh vực giao thông cũng không tránh khỏi những ảnh hưởng, nổi bật là tình trạng ùn tắc vào các giờ cao điểm tại các giao lộ, nơi mà phương tiện đi lại rất nhiều, làn đường hẹp cùng với việc phải chờ đèn đỏ quá lâu. Bài toán điều khiển tín hiệu giao thông tự động (TSC) được ra đời với mong muốn tìm ra các chiến lược thay đổi đèn tín hiệu tối ưu để điều hòa lưu lượng phương tiện tại các nút giao lớn.

Các phương pháp truyền thống như hệ thống hẹn giờ trước hay hệ thống điều khiển dựa trên quan sát thực tế của con người mang lại hiệu quả nhất định, nhưng vẫn chưa thích ứng tốt với nhiều trạng thái giao thông khác nhau và đòi hỏi chi phí cao cũng như phụ thuộc vào nhân sự vận hành. Vậy liệu có giải pháp nào vừa hiệu quả vừa giúp tiết kiệm chi phí và nguồn lực hơn ?

Trong những năm gần đây, học tăng cường (Reinforcement Learning - RL) đã thu hút nhiều sự quan tâm từ các nhà nghiên cứu với mục tiêu xây dựng và phát triển các hệ thống vận hành tự động, không cần sự can thiệp của con người. Với sự tiến bộ và cải tiến liên tục, RL đang dần cho thấy sự đa dạng và mạnh mẽ trong việc giải quyết các tác vụ phức tạp. Vậy việc áp dụng các thuật toán RL để tự động học và tìm ra các chiến lược cho một vấn đề cần tự động để giảm bớt sự phụ thuộc vào con người và tối ưu hiệu suất liệu có phải là một giải pháp hiệu quả hay không ?

Trong đề tài này, chúng tôi nghiên cứu về việc áp dụng một số thuật toán RL cho việc lựa chọn hành động và xây dựng các chiến lược điều khiển đèn tín hiệu cho bài toán Điều khiển tín hiệu giao thông tự động, bên cạnh đó các chiến lược được tìm ra vẫn phải đảm bảo đúng những nguyên tắc giao thông được quy định. Cụ thể:

- *Đầu vào*: trạng thái của lưu lượng giao thông hiện tại
- *Đầu ra*: trạng thái của lưu lượng giao thông được cho phép tại bước thời gian tiếp theo.

## MỤC TIÊU

*(Viết trong vòng 3 mục tiêu, lưu ý về tính khả thi và có thể đánh giá được)*

Trong đề tài này, chúng tôi hướng đến việc hoàn thành hai mục tiêu chính, bao gồm:

- Áp dụng các thuật toán Học tăng cường để tìm ra các chiến lược cho bài toán Điều khiển tín hiệu thông tự động với môi trường là các bản đồ mô phỏng có trong ứng dụng mô phỏng các tình huống giao thông SUMO
- So sánh, đánh giá độ hiệu quả của các chiến lược tìm lược bởi các thuật toán Học tăng cường với các chiến lược được thiết kế dựa quan sát và các chiến lược được thiết kế ngẫu nhiên.

## NỘI DUNG VÀ PHƯƠNG PHÁP

*(Viết nội dung và phương pháp thực hiện để đạt được các mục tiêu đã nêu)*

- Nội dung:
  - Tìm hiểu bài toán Điều khiển tín hiệu giao thông tự động và những hướng tiếp cận đã có trước đó để giải quyết bài toán.
  - Tìm hiểu về các thuật toán Học tăng cường.
  - Tìm hiểu về các môi trường mô phỏng các trạng thái giao thông.
  - Mô hình hóa bài toán Điều khiển tín hiệu giao thông tự động dựa trên bộ mô phỏng giao thông được đề xuất.
  - Chạy thực nghiệm và đánh giá hiệu suất của các chiến lược tìm được dựa trên bài toán đặt ra.
- Phương pháp:
  - Cài đặt ứng dụng SUMO - tạo ra các bản đồ mô phỏng giao thông bao gồm: i) Bản đồ đơn giao lộ; ii) Bản đồ có 7 giao lộ nằm dọc theo một

tuyến đường huyết mạch; iii) Bản đồ có nhiều giao lộ trong một khu vực.

- Thiết kế tác nhân để học ra các chiến lược ngẫu nhiên và một tác nhân khác học ra chiến lược dựa trên các quy tắc được định nghĩa trước (chiến lược mặc định).
- Cài đặt các thuật toán học tăng cường gồm: Deep Q-Network (DQN), Proximal Policy Optimization (PPO), Double DQN và một biến thể khác là Max Pressure Light (MP Light), cụ thể là thay đổi mô hình mạng huấn luyện.
- Huấn luyện các bộ điều khiển được thiết kế trên từng tình huống giao thông mô phỏng.
- Đánh giá các chiến lược tìm được với nhau, dựa trên các hàm điểm thưởng.

## KẾT QUẢ MONG ĐỢI

*(Viết kết quả phù hợp với mục tiêu đặt ra, trên cơ sở nội dung nghiên cứu ở trên)*

- Báo cáo các kỹ thuật Học tăng cường được áp dụng cho bài toán Điều khiển tín hiệu giao thông.
- Trình bày về các hàm điểm thưởng (chỉ số đánh giá ) để đánh giá hiệu suất của các chiến lược được tìm ra.
- Trình bày kết quả thực nghiệm, so sánh, đánh giá hiệu suất của các chiến lược tìm được trên 8 tình huống giao thông khác nhau.
- Áp dụng một số chiến lược tìm được lên bản mô phỏng tương ứng để minh họa thực tế hiệu suất của các chiến lược này.

## **TÀI LIỆU THAM KHẢO** (*Định dạng DBLP*)

- [1] Chacha Chen et al. “Toward A Thousand Lights: Decentralized Deep Reinforcement Learning for Large-Scale Traffic Signal Control”. In: AAAI Conference on Artificial Intelligence. 2020
- [2] Yit Kwong Chin et al. “Exploring Q-Learning Optimization in Traffic Signal Timing Plan Management”. In: 2011 Third International Conference on Computational Intelligence, Communication Systems and Networks. 2011, pp. 269274. DOI: 10.1109/CICSyN.2011.64
- [3] Guanjie Zheng et al. Learning Phase Competition for Traffic Signal Control. 2019. DOI: 10.48550/ARXIV.1905.04722.
- [4] Huichu Zhang et al. “CityFlow: A Multi-Agent Reinforcement Learning Environment for Large Scale City Traffic Scenario”. In: The World Wide Web Conference. ACM, 2019.
- [5] Liben Huang and Xiaohui Qu. “Improving traffic signal control operations using proximal policy optimization”. In: IET Intelligent Transport Systems (Oct. 2022)