BÁO CÁO ĐỒ ÁN TRAFIC SIGNAL CONTROL

Nguyễn Duy Hưng – 21520894

Thân Nguyễn Đức Hải – 21520810

# Giới thiệu đề tài

Trong những năm gần đây, vấn đề ùn tắc giao thông đã trở thành một thách thức lớn đối với các thành phố lớn trên toàn thế giới. Với sự gia tăng về số lượng phương tiện cùng với cơ sở hạ tầng giao thông còn hạn chế, việc quản lý giao thông hiệu quả và giảm thiểu ùn tắc trở thành một vấn đề cấp bách.

Trong lĩnh vực này, Traffic Signal Control là một giải pháp quản lý giao thông hiệu quả và đang được sử dụng phổ biến. Tuy nhiên, các phương pháp truyền thống của Traffic Signal Control vẫn còn rất hạn chế và không thể đáp ứng được yêu cầu của một hệ thống giao thông hiện đại.

Trong bối cảnh đó, đề tài này tập trung nghiên cứu và triển khai một hệ thống Traffic Signal Control mới sử dụng Reinforcement Learning (RL), một phương pháp học tăng cường trong Machine Learning. Hệ thống này sẽ được thiết kế để tự động điều khiển tín hiệu giao thông, giảm thiểu ùn tắc và tối ưu hóa luồng giao thông.

Mục tiêu của đồ án là xây dựng một hệ thống Traffic Signal Control mới, sử dụng RL để điều khiển tín hiệu giao thông, đáp ứng được các yêu cầu về hiệu quả và độ tin cậy. Báo cáo này sẽ trình bày chi tiết về kiến trúc hệ thống, phương pháp áp dụng RL vào Traffic Signal Control, kết quả thực hiện và đánh giá của hệ thống.

# II. Tổng quan về Traffic Signal Control và Reinforcement Learning

## A. Traffic Signal Control

Traffic Signal Control (TSC) là một phương pháp quản lý giao thông được sử dụng phổ biến, nhằm điều chỉnh tín hiệu đèn giao thông và tối ưu hóa luồng giao thông. Tuy nhiên, các phương pháp truyền thống của TSC vẫn chưa đáp ứng được yêu cầu của một hệ thống giao thông hiện đại. Vì vậy, nhiều nghiên cứu đã tìm cách áp dụng các phương pháp học máy để giải quyết vấn đề này.

## B. Reinforcement Learning

Reinforcement Learning (RL) là một phương pháp học tăng cường trong Machine Learning, trong đó một hệ thống phải học từ kinh nghiệm để đưa ra các quyết định tối ưu trong một môi trường động. RL được áp dụng rộng rãi trong nhiều lĩnh vực, bao gồm cả quản lý giao thông.

## C. Mô phỏng trạng thái giao thông

Trong đề tài này, chúng tôi sử dụng phương pháp mô phỏng trạng thái giao thông bằng Simulating Urban Mobility (SUMO), một phần mềm mô phỏng giao thông tiên tiến. SUMO cung cấp một môi trường mô phỏng thực tế để thử nghiệm các giải pháp quản lý giao thông, cho phép đánh giá hiệu quả của hệ thống trước khi triển khai thực tế.

## D. Học tăng cường

Chúng tôi áp dụng hai phương pháp học tăng cường là Q-learning và Deep Q-learning để phát triển hệ thống TSC mới. Q-learning là một thuật toán RL đơn giản, dựa trên bảng Q-value, để học cách đưa ra quyết định tối ưu trong một môi trường động. Deep Q-learning là một phiên bản của Q-learning sử dụng mạng nơ-ron sâu để xử lý các trạng thái phức tạp hơn. Deep Q-learning cho phép hệ thống học được các chiến lược tối ưu và cải thiện hiệu quả của TSC.

Do không thể sắp xếp đủ thời gian nên đề tài này xin phép chỉ chạy single agent, muti-agent trên q-learning và chỉ single agent trên deep neural network.

Trong phần này, chúng tôi đã giới thiệu về Traffic Signal Control, Reinforcement Learning và phương pháp mô phỏng trạng thái giao thông bằng SUMO. Chúng tôi cũng đã giới thiệu hai phương pháp học tăng cường là Q-learning và Deep Q-learning, được áp dụng để phát triển hệ thống TSC mới. Trong phần tiếp theo, chúng tôi sẽ trình bày chi tiết về kiến trúc hệ thống và phương pháp áp dụng RL vào TSC.

# III. Thiết kế và triển khai hệ thống

## A. Kiến trúc hệ thống

Hệ thống TSC mới được thiết kế dựa trên mô hình học tăng cường Q-learning và Deep Q-learning. Hệ thống được chia thành hai phần chính: mô phỏng SUMO và học tăng cường.

Mô phỏng SUMO: Hệ thống sử dụng phần mềm mô phỏng SUMO để tạo ra môi trường giao thông thực tế, bao gồm các tuyến đường, điểm giao cắt và các phương tiện di chuyển. Dữ liệu giao thông được thu thập từ mô phỏng và được sử dụng để đưa ra quyết định điều khiển tín hiệu giao thông.

Học tăng cường: Hệ thống sử dụng thuật toán Q-learning và Deep Q-learning để học cách điều khiển tín hiệu giao thông tối ưu. Hệ thống sử dụng mô hình Deep Q-learning để xử lý các trạng thái phức tạp của giao thông và đưa ra các quyết định tối ưu.

Mô hình hoá bài toán:

Agent: Mỗi agent được xác định là một nút giao thông.

Không gian trạng thái: không gian trạng thái được xác định bởi

* phase\_one\_hot là một vectơ được mã hóa một lần cho biết pha màu xanh lá cây đang hoạt động hiện tại.
* min\_green là một biến nhị phân cho biết min\_green giây đã trôi qua trong giai đoạn hiện tại chưa. Mục đích là kiểm tra có thể được phép chuyển đèn chưa.
* lane\_i\_density là số lượng xe trong làn đi vào i chia cho tổng sức chứa của làn.
* lane\_i\_queue là số lượng xe xếp hàng (tốc độ dưới 0,1 m/s) trong làn đường đến i chia cho tổng sức chứa của làn đường.

Hành động: là việc lựa chọn hành động và điều chỉnh thời gian của hành động được xuất hiện phía sau. Tuỳ bản đồ chúng ta có những actions space khác nhau. Ví dụ, trong bản đồ 2way-single-intersection chúng ta có 4 hành động được điều chỉnh luận phiên và nhiệm vụ của hệ thống là điều chỉnh thời gian của các hành động để cải thiện các hệ số để đạt được reward tối đa.

A close-up of a grid

Description automatically generated

Hình 1: 4 hành động trong ngã tư giao nhau 2 chiều.

Phần thưởng: chúng tôi dựa vào giá trị để xác định giá trị của trạng thái:

* "diff-waiting-time": phần thưởng là tổng độ trễ (tổng thời gian chờ của tất cả các phương tiện đang đến gần) đã thay đổi như thế nào so với bước thời gian trước đó.

B. Phương pháp áp dụng RL vào TSC

Để áp dụng RL vào TSC, chúng tôi sử dụng mô hình học tăng cường Q-learning và Deep Q-learning. Hệ thống được huấn luyện trên dữ liệu giao thông được thu thập từ mô phỏng SUMO để tối ưu hóa đường đi và giảm thiểu thời gian chờ đợi của các phương tiện.

## C. Bộ dữ liệu:

Bộ dữ liệu: Để huấn luyện hệ thống, chúng tôi sử dụng mô phỏng SUMO để tạo ra một môi trường giao thông thực tế. Dữ liệu giao thông được thu thập từ mô phỏng này và được sử dụng để đào tạo hệ thống. Ở đây chúng tôi sẽ thực hiện chạy trên 4 bản đồ:

A cross-section of a cross-section

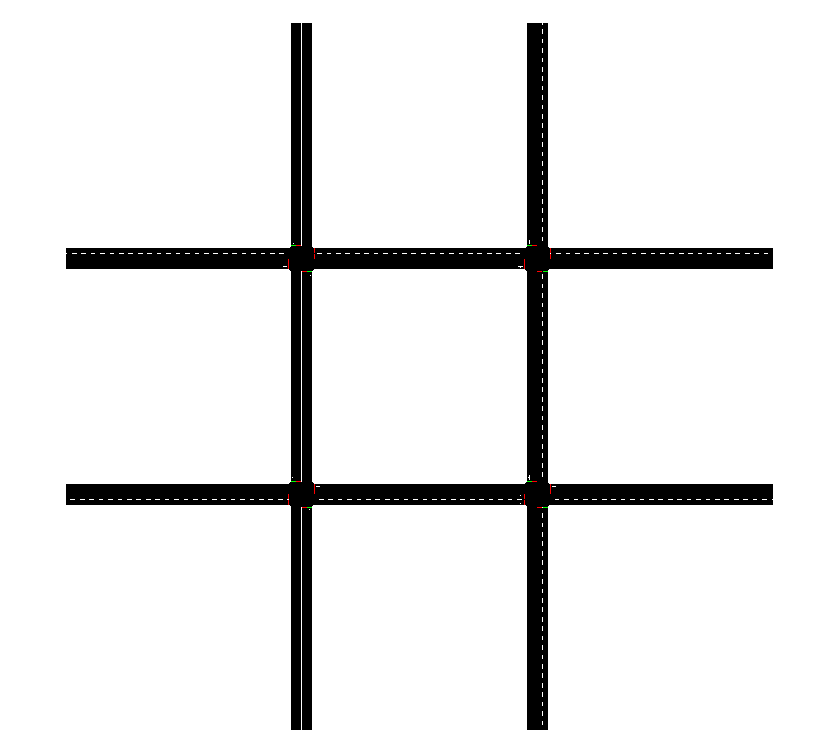
Description automatically generated

Hình 2: bản đồ single-intersection

A road with a cross in the middle

Description automatically generated

Hình 3: Bản đồ 2way-single-intersection



Hình 4: Bản đồ 2x2grid

A black antenna with green lights

Description automatically generated

Hình 5: bản đồ resco-cologne1

Trong phần này, chúng tôi đã giới thiệu về kiến trúc hệ thống, phương pháp áp dụng RL vào TSC và các tham số cần thiết cho quá trình huấn luyện. Chúng tôi cũng đã mô tả về quá trình kiểm thử và đánh giá hiệu quả của hệ thống. Trong phần tiếp theo, chúng tôi sẽ trình bày chi tiết về kết quả và đánh giá của hệ thống.

# IV. Kết quả và đánh giá

Sau khi thực hiện chạy thử nghiệm các phương pháp học tăng cường nêu trên, ta có kết quả:

A graph with blue lines

Description automatically generatedA graph with blue lines

Description automatically generated

Hình 6 và 7: Kết quả chạy thử(bên trái) và học (bên phải) của bản đồ 2way-single-intersection Q-learning.

A graph with blue lines

Description automatically generatedA graph showing steps and steps

Description automatically generated

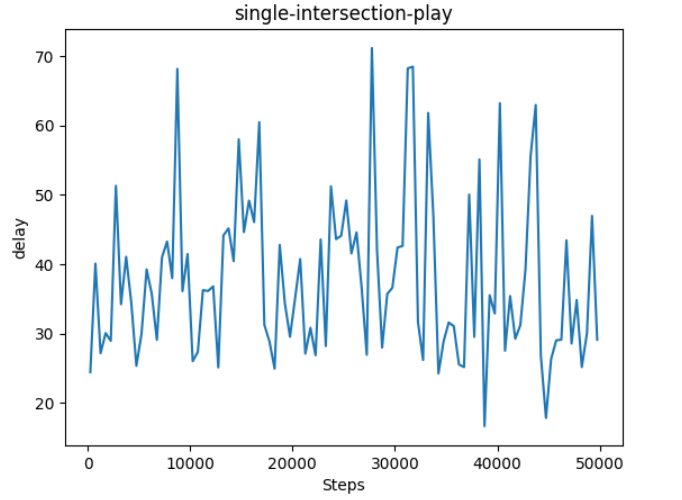
Hình 8 và 9: Kết quả chạy thử(bên trái) và học (bên phải) của bản đồ 2x2grid Q-learning.

A graph of steps and steps

Description automatically generatedA graph of steps and steps

Description automatically generated

Hình 10 và 11: Kết quả chạy thử(bên trái) và học (bên phải) của bản đồ resco-cologne1 Q-learning.

A graph with blue lines

Description automatically generated

Hình 12 và 13: Kết quả chạy thử(bên trái) và học (bên phải) của bản đồ single-intersection Q-learning.

A graph of steps

Description automatically generatedA graph with blue lines

Description automatically generatedA graph with blue lines

Description automatically generated

Hình 14,15,16: Lần lượt là kết quả khi học của deep-Q-learning khi thực hiện chạy 3 bản đồ resco-cologne1, single-intersection và 2way-single-intersection.

**Đánh giá:**

Sau khi thực hiện cả 2 thuật toán thì ta nhận thấy rằng, cả 2 thuật toán đều có khả năng thích ứng với môi trường và cải thiện được hiệu suất đến một mức nhất định nào đó.

So sánh về hiệu suất thì dqn cho một hiệu suất học tập tốt hơn khá nhiều so với q-learning, khả năng thích nghi, thời gian chạy, khả năng khám phá của deep q-learning là tốt. Dù vậy lượng dữ liệu để deep q-learning cần để đặt được chiến lược tối ưu là tương đối lớn.

Ngoài ra, deep q-learning mang lại tính ổn định tương đối tốt, tránh các trường hợp overfitting. Q-learning có thể dễ dàng bị ảnh hưởng bởi các động thái ngẫu nhiên của môi trường, dẫn đến sai lệch giá trị Q.

Tóm lại, hệ thống áp dụng RL vào TSC là khả thi. Tuy nhiên, để đạt được hiệu quả tối ưu, cần phải có sự cân nhắc kỹ lưỡng trong việc lựa chọn thuật toán, cấu trúc mô hình, tham số của mô hình và đáp ứng các yêu cầu của người dùng và môi trường.

# V. Kết luận và hướng phát triển

## Kết luận

* Mục tiêu của đồ án là tìm hiểu, nghiên cứu, ứng dụng học tăng cường (RL) vào điều khiển tín hiệu đèn giao thông. Từ những thông tin ở phía trên, ta có thể nhận thấy tính ứng dụng cũng như tính khả thi của mục tiêu đang hướng đến thông qua 2 phương pháp là Q-learning và deep-Q-learning. Kết quả của đồ án là một trải nghiệm giúp chúng tôi hiểu hơn về RL cũng như ứng dụng của nó trong các lĩnh vực thực tế trong đời sống. Kết quả này còn mở ra rất nhiều hướng phát triển của đồ án trong tương lai như ứng dụng các phương pháp học tăng cường khác như MPLight, FMA2C,… hay ứng dụng các phương pháp học tăng cường vào các lĩnh vực khác trong đời sống,…
* Do giới hạn về thời gian, kiến thức cũng như điều kiện để thực hiện dự án nên còn kết quả trên còn nhiều sai sót cũng như chưa thể tìm ra một cái nhìn toàn diện về đề tài này. Tuy vậy, các kết quả và nhận xét trong báo cáo cũng cung cấp cho chúng ta một bức tranh về tiềm năng và thách thức trong việc áp dụng RL vào các lĩnh vực thực tế nói chung và traffic signal control nói riêng.

# VI. Tài liệu tham khảo

L.N.Alegre.Sumo-rl.https://github.com/LucasAlegre/sumo-rl,2019.

JamesAult and GuniSharon. Reinforcement Learning Benchmarks for Traffic Signal Control