

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC ĐẠI NAM
KHOA CÔNG NGHỆ THÔNG TIN



BÁO CÁO BÀI TẬP LỚN

SẮP XẾP THỜI KHÓA BIỂU

Sinh viên thực hiện : Tạ Văn Phong
Nguyễn Phương Nam
Nguyễn Phương Nam
Mai Ngọc Vui

Ngành : Công nghệ thông tin

Giảng viên hướng dẫn : ThS. Lê Thị Thùy Trang

Mục lục

| | | |
|----------|--|-----------|
| 1 | TỔNG QUAN VỀ BÀI TOÁN LẬP LỊCH HỌC | 1 |
| 1.1 | Giới thiệu và lý do chọn đề tài | 1 |
| 1.2 | Mục tiêu và phạm vi nghiên cứu | 2 |
| 1.2.1 | Mục tiêu nghiên cứu | 2 |
| 1.2.2 | Phạm vi nghiên cứu | 3 |
| 1.3 | Độ phức tạp của bài toán lập lịch học | 4 |
| 1.3.1 | Không gian tìm kiếm lớn | 4 |
| 1.3.2 | Xung đột thời gian | 5 |
| 1.3.3 | Sự kết hợp giữa ràng buộc cứng và ràng buộc mềm | 5 |
| 1.3.4 | Tối ưu hóa nhiều tiêu chí cùng lúc | 6 |
| 1.3.5 | Bài toán thuộc nhóm NP-hard | 6 |
| 1.4 | Các phương pháp tiếp cận bài toán | 7 |
| 1.4.1 | Phương pháp truyền thống | 7 |
| 1.4.2 | Phương pháp heuristic (Phương pháp tìm kiếm gần đúng) | 8 |
| 1.4.3 | Phương pháp dựa trên thuật toán tiến hóa | 9 |
| 1.4.4 | Phương pháp học máy (Machine Learning) | 10 |
| 2 | Phân tích yêu cầu bài toán | 12 |
| 2.1 | Mô tả bài toán | 12 |
| 2.1.1 | Các ràng buộc | 12 |
| 2.1.2 | Mục tiêu của bài toán | 13 |
| 2.2 | Mô tả thuật toán sử dụng | 14 |
| 2.2.1 | Giới thiệu thuật toán di truyền (Genetic Algorithm - GA) | 14 |
| 2.2.2 | Áp dụng thuật toán di truyền vào bài toán lập lịch | 14 |
| 2.2.3 | Kết luận | 16 |
| 2.3 | Phân tích mã nguồn thuật toán tạo lịch học tự động bằng thuật toán di truyền | 17 |
| 2.3.1 | Mô tả tổng quan về chương trình | 17 |
| 2.3.2 | Các lớp đối tượng chính trong chương trình | 17 |
| 2.3.3 | Cách hoạt động của thuật toán di truyền trong chương trình | 18 |

| | | |
|----------|---|-----------|
| 2.3.4 | Khởi tạo dữ liệu đầu vào | 19 |
| 2.3.5 | Chạy thuật toán và hiển thị kết quả | 20 |
| 3 | Kết luận | 21 |
| 3.1 | Đánh giá chung về đề tài | 21 |
| 3.1.1 | Những kết quả đạt được | 21 |
| 3.1.2 | Một số hạn chế của đề tài | 22 |
| 3.2 | Hướng phát triển của đề tài | 22 |
| 3.3 | Kết luận | 23 |

Chương 1

TỔNG QUAN VỀ BÀI TOÁN LẬP LỊCH HỌC

1.1 Giới thiệu và lý do chọn đề tài

Trong môi trường giáo dục, việc lập lịch học là một bài toán quan trọng và có ảnh hưởng trực tiếp đến hoạt động giảng dạy và học tập của giáo viên, sinh viên. Một thời khóa biểu hợp lý không chỉ giúp tối ưu hóa việc sử dụng phòng học, thiết bị giảng dạy mà còn giảm thiểu xung đột giữa các lớp học, đảm bảo sự phân bố đồng đều thời gian học cho sinh viên và giảng viên.

Tuy nhiên, việc xây dựng một thời khóa biểu hợp lý không phải là một nhiệm vụ đơn giản. Quá trình lập lịch học đòi hỏi phải cân nhắc nhiều yếu tố như:

- Số lượng giảng viên và thời gian giảng dạy của họ.
- Số lượng sinh viên và các lớp học.
- Số lượng phòng học và cơ sở vật chất hỗ trợ.
- Ràng buộc về thời gian của từng môn học.
- Các yêu cầu đặc biệt như môn học thực hành, lý thuyết, nhóm học riêng biệt.

Bài toán lập lịch học thuộc nhóm bài toán NP-hard, có nghĩa là không thể tìm ra lời giải tối ưu bằng phương pháp thử tất cả các phương án trong thời gian ngắn. Do đó, các phương pháp tìm kiếm tối ưu như thuật toán di truyền (Genetic Algorithm - GA) được sử dụng để đưa ra phương án tốt nhất trong một khoảng thời gian hợp lý.

Với tầm quan trọng của bài toán và những thách thức trong việc tìm kiếm lời giải tối ưu, nhóm nghiên cứu lựa chọn đề tài "Ứng dụng thuật toán di truyền trong bài toán lập lịch học",

nhằm xây dựng một hệ thống tự động hóa quá trình lập lịch, giúp giảm thiểu công sức và thời gian cho các nhà quản lý giáo dục.

1.2 Mục tiêu và phạm vi nghiên cứu

1.2.1 Mục tiêu nghiên cứu

Bài toán lập lịch học là một trong những bài toán tối ưu hóa quan trọng trong lĩnh vực giáo dục, ảnh hưởng trực tiếp đến chất lượng giảng dạy và trải nghiệm học tập của sinh viên. Việc xây dựng một thời khóa biểu hợp lý đòi hỏi phải xem xét nhiều ràng buộc phức tạp như số lượng giảng viên, phòng học, giới hạn thời gian, cũng như sự phân bố hợp lý giữa các môn học.

Để giải quyết bài toán này, đề tài đặt ra các mục tiêu nghiên cứu cụ thể như sau:

- Phát triển một hệ thống lập lịch học tự động sử dụng thuật toán di truyền (GA)
 - Xây dựng một mô hình có khả năng tự động tạo ra thời khóa biểu hợp lệ dựa trên các ràng buộc đầu vào.
 - Đảm bảo hệ thống có thể tạo ra lịch học tối ưu trong thời gian ngắn, giảm bớt công sức của người quản lý giáo dục.
 - Cải thiện khả năng ứng dụng thực tế của hệ thống bằng cách cung cấp giao diện trực quan giúp người dùng dễ dàng tương tác và điều chỉnh lịch học nếu cần.
- Cải thiện chất lượng thời khóa biểu thông qua tối ưu hóa các ràng buộc
 - Hạn chế tối đa xung đột lịch học (ví dụ: một giảng viên không thể dạy hai lớp cùng một thời điểm, một phòng học không thể được sử dụng đồng thời cho hai lớp).
 - Tối ưu hóa việc sử dụng phòng học để tránh tình trạng phòng học bị bỏ trống hoặc quá tải.
 - Đảm bảo thời gian giảng dạy hợp lý, tránh trường hợp giảng viên hoặc sinh viên có lịch học quá dày đặc trong một ngày.
 - Phân bổ thời gian học hợp lý giữa các buổi sáng, chiều và tối nhằm cân bằng khối lượng giảng dạy.
- Đánh giá hiệu suất của thuật toán di truyền trong bài toán lập lịch học
 - Thực hiện các thử nghiệm với các bộ dữ liệu khác nhau để kiểm tra tính hiệu quả của thuật toán di truyền.

- So sánh thuật toán di truyền với các phương pháp truyền thống như lập lịch thủ công hoặc thuật toán tìm kiếm vét cạn.
- Phân tích độ chính xác, tốc độ xử lý và khả năng mở rộng của hệ thống khi số lượng khóa học, giảng viên và phòng học tăng lên.

Bằng cách đạt được các mục tiêu trên, đề tài hướng tới việc tạo ra một công cụ hỗ trợ đắc lực cho công tác quản lý giáo dục, giúp giảm bớt gánh nặng cho người lập lịch và tối ưu hóa việc phân bổ tài nguyên giảng dạy.

1.2.2 Phạm vi nghiên cứu

Bài toán lập lịch học có thể có nhiều biến thể và mức độ phức tạp khác nhau. Để đảm bảo tính khả thi của đề tài trong khuôn khổ một bài tập lớn, phạm vi nghiên cứu được giới hạn như sau:

Đối tượng và dữ liệu đầu vào

- Đề tài tập trung vào lập lịch học cho một trường đại học hoặc cao đẳng, trong đó có nhiều lớp học, giảng viên và phòng học.
- Dữ liệu đầu vào bao gồm danh sách giảng viên, môn học, lớp học, phòng học và các ràng buộc liên quan.
- Giả định rằng tất cả dữ liệu đầu vào đều có sẵn và không thay đổi trong quá trình lập lịch (không có giảng viên hay lớp học phát sinh sau khi hệ thống đã tạo xong thời khóa biểu).

Các ràng buộc được xem xét

- Ràng buộc cứng (bắt buộc phải thỏa mãn):
 - Mỗi giảng viên chỉ có thể giảng dạy một lớp tại một thời điểm.
 - Mỗi phòng học chỉ có thể chứa một lớp học tại một thời điểm.
 - Giảng viên chỉ có thể giảng dạy các môn học thuộc chuyên môn của mình.
 - Một môn học có thể có nhiều lớp khác nhau, nhưng không được trùng giờ nếu cùng một giảng viên giảng dạy.
 - Mỗi lớp phải có đủ số tiết học theo yêu cầu của môn học.
- Ràng buộc mềm (ưu tiên tối ưu hóa nhưng có thể vi phạm nếu cần thiết):

- Phân bổ lịch học hợp lý để giảm thiểu thời gian trống giữa các tiết học của giảng viên và sinh viên.
- Hạn chế số lượng tiết học liên tiếp của một giảng viên hoặc sinh viên trong một ngày để tránh quá tải.
- Cân bằng giữa các buổi học sáng, chiều và tối để tối ưu hóa việc sử dụng phòng học.

Giới hạn và loại trừ

- Không xét đến các yếu tố phức tạp như:
 - Thay đổi lịch học đột xuất do giảng viên vắng mặt hoặc sự kiện đặc biệt.
 - Lịch học thay đổi theo tuần (ví dụ: một môn học có thể có lịch học khác nhau mỗi tuần).
 - Sự ưu tiên cá nhân của giảng viên hoặc sinh viên (ví dụ: giảng viên thích dạy vào buổi sáng hơn buổi chiều).
- Không tích hợp các phương pháp học máy phức tạp như học tăng cường (Reinforcement Learning) hoặc mạng nơ-ron nhân tạo (Artificial Neural Networks).

Với phạm vi nghiên cứu này, đề tài đảm bảo tính khả thi trong việc phát triển và thử nghiệm mô hình, đồng thời vẫn giữ được tính thực tiễn khi áp dụng vào các hệ thống lập lịch học tự động.

1.3 Độ phức tạp của bài toán lập lịch học

Bài toán lập lịch học là một trong những bài toán tối ưu hóa phức tạp, đòi hỏi việc sắp xếp lịch học sao cho thỏa mãn đồng thời nhiều ràng buộc khác nhau, cả cứng (bắt buộc) và mềm (ưu tiên). Điều này làm cho bài toán có độ phức tạp cao, đặc biệt khi số lượng lớp học, giảng viên và phòng học tăng lên.

Dưới đây là một số thách thức chính khiến bài toán lập lịch học trở nên khó khăn:

1.3.1 Không gian tìm kiếm lớn

Bài toán lập lịch học có số lượng tổ hợp lịch học cực kỳ lớn do số lượng biến số cần xử lý tăng nhanh theo quy mô của hệ thống. Chẳng hạn, giả sử có:

- N lớp học
- M giảng viên

- P phòng học
- T khoảng thời gian trong tuần

Tổng số cách lập lịch có thể có là:

$$(Slphc) \times (Sgingvinphhp) \times (Sphngkhdng) \times (Skhongthigian)$$

Khi số lượng lớp học, giảng viên và phòng học tăng lên, số lượng tổ hợp có thể tăng lên theo cấp số nhân, dẫn đến một không gian tìm kiếm khổng lồ.

Vì vậy, việc tìm kiếm một lịch học tối ưu trong một không gian lớn như vậy bằng cách duyệt toàn bộ các khả năng là không khả thi trong thực tế.

1.3.2 Xung đột thời gian

Một trong những thách thức lớn nhất khi lập lịch học là xử lý các xung đột về thời gian, bao gồm:

- Xung đột giảng viên: Một giảng viên không thể dạy hai lớp cùng một thời điểm.
- Xung đột phòng học: Một phòng học không thể được sử dụng cho hai lớp học cùng một lúc.
- Xung đột môn học: Một sinh viên không thể có hai môn học trùng giờ nhau.

Nếu số lượng lớp học và giảng viên lớn, việc kiểm tra và giải quyết xung đột sẽ trở nên rất khó khăn.

1.3.3 Sự kết hợp giữa ràng buộc cứng và ràng buộc mềm

Bài toán lập lịch học không chỉ yêu cầu tìm ra một lịch học hợp lệ (thỏa mãn các ràng buộc cứng) mà còn cần tối ưu hóa theo các ràng buộc mềm.

- Ràng buộc cứng (bắt buộc phải thỏa mãn):
 - Mỗi lớp phải có đủ số tiết học theo yêu cầu.
 - Mỗi giảng viên chỉ có thể dạy một lớp tại một thời điểm.
 - Mỗi phòng học chỉ có thể chứa một lớp tại một thời điểm.
 - Một lớp chỉ có thể học một môn duy nhất trong một khoảng thời gian.
- Ràng buộc mềm (có thể vi phạm nhưng nên hạn chế):

- Giảm thiểu khoảng trống giữa các tiết học để tối ưu thời gian giảng dạy của giảng viên và sinh viên.
- Hạn chế việc dồn quá nhiều lớp học của một giảng viên vào một buổi.
- Tối ưu hóa việc sử dụng phòng học để tránh phòng bị bỏ trống hoặc sử dụng không hợp lý.

Sự kết hợp giữa các ràng buộc cứng và mềm làm cho bài toán trở nên phức tạp hơn, vì việc thỏa mãn tất cả ràng buộc cùng lúc là rất khó, và cần có sự đánh đổi giữa các tiêu chí tối ưu hóa.

1.3.4 Tối ưu hóa nhiều tiêu chí cùng lúc

Bài toán lập lịch học không chỉ yêu cầu tìm ra một lịch học hợp lệ mà còn cần tối ưu hóa theo nhiều tiêu chí khác nhau:

- Công bằng trong phân bố lịch giảng dạy: Tránh tình trạng một giảng viên có quá nhiều tiết dạy trong một ngày, trong khi giảng viên khác có lịch học rải rác không đều.
- Tận dụng hiệu quả phòng học: Giảm thiểu số phòng học bị bỏ trống hoặc tình trạng phòng quá tải.
- Hạn chế thời gian trống giữa các tiết học: Giúp giảng viên và sinh viên không phải chờ đợi quá lâu giữa các tiết học.
- Phân bố lịch học hợp lý theo tuần: Đảm bảo rằng số tiết học của một lớp không bị dồn quá nhiều vào một ngày mà được phân bố đồng đều trong tuần.

Do có nhiều tiêu chí cần tối ưu, việc tìm ra một lịch học tốt nhất đòi hỏi sự cân bằng giữa các yếu tố trên, khiến bài toán trở nên phức tạp hơn.

1.3.5 Bài toán thuộc nhóm NP-hard

Bài toán lập lịch học được chứng minh là một bài toán NP-hard, có nghĩa là:

- Không có thuật toán giải quyết bài toán một cách tối ưu trong thời gian đa thức (polynomial time).
- Phương pháp duyệt toàn bộ không gian tìm kiếm (brute force) là không khả thi do số lượng tổ hợp quá lớn.
- Cần sử dụng các phương pháp tìm kiếm gần đúng (heuristic) để tìm ra lời giải chấp nhận được trong thời gian hợp lý.

Một số phương pháp được sử dụng để giải quyết bài toán này bao gồm:

- Thuật toán di truyền (Genetic Algorithm - GA): Giả lập quá trình tiến hóa tự nhiên để tìm kiếm lời giải tối ưu.
- Thuật toán tìm kiếm Tabu (Tabu Search): Sử dụng danh sách cấm để tránh lặp lại các lời giải kém hiệu quả.
- Thuật toán mô phỏng luyện kim (Simulated Annealing - SA): Bắt chước quá trình nung chảy và làm lạnh của kim loại để tìm kiếm lời giải tốt hơn.

Trong đề tài này, thuật toán di truyền (GA) được chọn làm phương pháp chính để giải quyết bài toán lập lịch học, nhờ khả năng tìm kiếm lời giải tốt trong không gian tìm kiếm lớn và sự linh hoạt trong việc xử lý các ràng buộc.

1.4 Các phương pháp tiếp cận bài toán

Bài toán lập lịch học là một bài toán tối ưu hóa phức tạp, do đó có nhiều phương pháp tiếp cận khác nhau để giải quyết nó. Tùy vào mức độ phức tạp và yêu cầu thực tế, mỗi phương pháp đều có những ưu và nhược điểm riêng. Dưới đây là các cách tiếp cận phổ biến:

1.4.1 Phương pháp truyền thống

Các phương pháp truyền thống chủ yếu dựa vào kinh nghiệm của người lập lịch hoặc sử dụng các thuật toán tìm kiếm đơn giản để tạo ra lịch học.

Lập lịch thủ công

- Cách thực hiện: Người lập lịch dựa vào kinh nghiệm và quy tắc sẵn có để phân công thời khóa biểu cho các lớp học, giảng viên và phòng học.
- Ưu điểm:
 - Có thể linh hoạt điều chỉnh dựa trên yêu cầu thực tế.
 - Phù hợp với quy mô nhỏ, số lượng môn học và giảng viên ít.
- Nhược điểm:
 - Mất nhiều thời gian, dễ xảy ra lỗi do chủ quan.
 - Khó tối ưu khi số lượng khóa học, giảng viên và phòng học tăng lên.
 - Không thể đảm bảo tìm được giải pháp tối ưu tốt nhất.

Sử dụng quy tắc và thuật toán tìm kiếm vét cạn (Brute Force Search)

- Cách thực hiện: Kiểm tra tất cả các tổ hợp lịch học có thể có để tìm ra phương án tối ưu nhất.
- Ưu điểm:
 - Đảm bảo tìm được phương án tối ưu nếu có thể duyệt hết các khả năng.
 - Dễ triển khai với quy mô nhỏ.
- Nhược điểm:
 - Độ phức tạp tính toán rất cao, thời gian xử lý tăng theo cấp số nhân khi số lượng lớp học và giảng viên tăng.
 - Không khả thi với hệ thống lớn do số lượng tổ hợp quá lớn.

Kết luận: Phương pháp truyền thống không phù hợp với bài toán lập lịch học trong thực tế vì không đảm bảo tính tối ưu và mất nhiều thời gian.

1.4.2 Phương pháp heuristic (Phương pháp tìm kiếm gần đúng)

Heuristic là các thuật toán tìm kiếm gần đúng, giúp tìm ra lời giải khả thi trong một không gian tìm kiếm lớn mà không cần phải kiểm tra toàn bộ các tổ hợp có thể.

Thuật toán tham lam (Greedy Algorithm)

- Cách thực hiện: Ở mỗi bước, thuật toán chọn giải pháp tốt nhất có thể tại thời điểm đó mà không quan tâm đến tương lai.
- Ưu điểm:
 - Tốc độ nhanh, đơn giản, dễ triển khai.
 - Hiệu quả khi bài toán có cấu trúc phù hợp với cách tiếp cận tham lam.
- Nhược điểm:
 - Không đảm bảo tìm được giải pháp tối ưu toàn cục.
 - Có thể mắc kẹt ở lời giải tối ưu cục bộ mà không tìm ra phương án tốt nhất.

Thuật toán tìm kiếm Tabu (Tabu Search - TS)

- Cách thực hiện: Dựa trên tìm kiếm cục bộ, nhưng có thêm cơ chế ghi nhớ các giải pháp đã xét trước đó để tránh quay lại những lời giải kém hiệu quả.
- Ưu điểm:
 - Tránh được bẫy tối ưu cục bộ.
 - Có khả năng tìm được lời giải tốt hơn so với thuật toán tham lam.
- Nhược điểm:
 - Cần tinh chỉnh các tham số để đạt hiệu quả cao.
 - Có thể tốn thời gian xử lý nếu không được tối ưu tốt.

Kết luận: Phương pháp heuristic giúp tìm lời giải tốt hơn so với phương pháp truyền thống nhưng vẫn chưa đảm bảo tìm ra phương án tối ưu nhất.

1.4.3 Phương pháp dựa trên thuật toán tiến hóa

Các thuật toán tiến hóa mô phỏng quá trình tiến hóa tự nhiên để tìm kiếm lời giải tối ưu bằng cách sử dụng các phép toán như chọn lọc, lai ghép và đột biến.

Thuật toán di truyền (Genetic Algorithm - GA)

- Cách thực hiện:
 1. Khởi tạo quần thể: Tạo ra một tập hợp các thời khóa biểu ban đầu.
 2. Đánh giá độ phù hợp (fitness function): Tính toán mức độ hợp lệ của mỗi cá thể (thời khóa biểu).
 3. Lựa chọn cá thể tốt nhất: Chọn những cá thể có độ phù hợp cao để sinh ra thế hệ mới.
 4. Lai ghép (Crossover): Kết hợp hai cá thể để tạo ra cá thể con có đặc điểm tốt từ cha mẹ.
 5. Đột biến (Mutation): Thay đổi ngẫu nhiên một phần lịch học để tăng tính đa dạng.
 6. Lặp lại: Quá trình này tiếp tục cho đến khi tìm được lời giải phù hợp.
- Ưu điểm:
 - Có khả năng tìm ra lời giải tốt trong không gian tìm kiếm lớn.

- Có thể xử lý nhiều ràng buộc phức tạp cùng lúc.
- Linh hoạt, có thể cải thiện chất lượng lời giải qua nhiều thế hệ.
- Nhược điểm:
 - Cần thời gian tính toán dài hơn so với phương pháp heuristic.
 - Hiệu suất phụ thuộc vào cách thiết lập tham số (số thế hệ, tỷ lệ lai ghép, tỷ lệ đột biến, v.v.).

Kết luận: Thuật toán di truyền là phương pháp được chọn trong đề tài này do khả năng tìm kiếm lời giải tối ưu trong không gian tìm kiếm lớn.

1.4.4 Phương pháp học máy (Machine Learning)

Các phương pháp học máy có thể được sử dụng để tối ưu hóa bài toán lập lịch học bằng cách học từ dữ liệu lịch sử và đưa ra quyết định dựa trên các mô hình dự đoán.

Học tăng cường (Reinforcement Learning - RL)

- Cách thực hiện:
 - Hệ thống học cách đưa ra quyết định tối ưu bằng cách nhận phản hồi (reward) khi đưa ra quyết định tốt và bị phạt (penalty) khi đưa ra quyết định không tốt.
 - Qua quá trình thử nghiệm và cải tiến, thuật toán học được chiến lược lập lịch tối ưu.
- Ưu điểm:
 - Có khả năng tự động học và thích ứng với dữ liệu mới.
 - Có thể tìm ra chiến lược lập lịch hiệu quả theo thời gian.
- Nhược điểm:
 - Cần lượng dữ liệu huấn luyện lớn.
 - Đòi hỏi nhiều tài nguyên tính toán.

Mạng nơ-ron nhân tạo (Artificial Neural Networks - ANN)

- Cách thực hiện:
 - Sử dụng dữ liệu lịch học trong quá khứ để huấn luyện mô hình dự đoán lịch học tối ưu.

- Mô hình có thể học được các mẫu tối ưu trong việc phân bổ lịch học.
- Ưu điểm:
 - Có khả năng tự động hóa và tối ưu hóa lịch học dựa trên dữ liệu thực tế.
- Nhược điểm:
 - Cần thời gian huấn luyện mô hình lâu.
 - Không dễ triển khai nếu không có tập dữ liệu đầy đủ.

Chương 2

Phân tích yêu cầu bài toán

2.1 Mô tả bài toán

Bài toán đặt ra là xây dựng một hệ thống tự động sắp xếp thời khóa biểu cho các lớp học dựa trên các ràng buộc cứng (bắt buộc) và mềm (ưu tiên). Hệ thống này phải đảm bảo việc phân bổ hợp lý tài nguyên giảng dạy (bao gồm phòng học, giảng viên, thời gian) nhằm tránh xung đột và tối ưu hóa chất lượng lịch học. Việc lập lịch cần đáp ứng các yêu cầu của nhà trường, giảng viên và sinh viên một cách hiệu quả nhất, đồng thời linh hoạt với những thay đổi trong dữ liệu đầu vào như số lượng lớp học, môn học, phòng học và giảng viên.

2.1.1 Các ràng buộc

Ràng buộc cứng (Bắt buộc)

Ràng buộc cứng là những điều kiện bắt buộc mà thời khóa biểu phải tuân thủ để đảm bảo tính hợp lệ của lịch học:

- Một giảng viên không thể dạy hai lớp cùng một thời điểm.
- Một lớp không thể học hai môn cùng một thời điểm.
- Một phòng học không thể được sử dụng cho nhiều lớp cùng lúc.
- Một môn học phải được xếp đúng số tiết quy định trong tuần.
- Các lớp học phải được phân bổ vào các khung giờ hợp lệ theo quy định của nhà trường.
- Các môn học có yêu cầu đặc biệt về phòng học (ví dụ: phòng thí nghiệm, phòng máy tính) phải được xếp vào đúng loại phòng phù hợp.
- Một số môn học có thể có điều kiện tiên quyết, yêu cầu phải được xếp sau các môn học

khác.

Ràng buộc mềm (Ưu tiên)

Ràng buộc mềm là những tiêu chí được ưu tiên để cải thiện chất lượng lịch học, nhưng không bắt buộc phải tuân thủ tuyệt đối:

- Tránh xếp quá nhiều tiết liên tiếp cho một giảng viên để tránh quá tải.
- Phân bổ đều lịch học trong tuần, tránh tập trung quá nhiều môn vào một ngày.
- Ưu tiên xếp lịch theo mong muốn của giảng viên (nếu có thể đáp ứng mà không vi phạm ràng buộc cứng).
- Giữ số lượng buổi học cân bằng giữa các lớp để tối ưu thời gian giảng dạy.
- Tránh xếp tiết học vào những khung giờ không thuận lợi (như tiết đầu buổi sáng hoặc tiết cuối buổi chiều) nếu có thể.
- Đảm bảo giảng viên không bị dạy quá nhiều lớp khác nhau trong cùng một ngày, giúp họ chuẩn bị bài giảng tốt hơn.

2.1.2 Mục tiêu của bài toán

Mục tiêu chính của bài toán là xây dựng một hệ thống tối ưu hóa thời khóa biểu dựa trên thuật toán di truyền (Genetic Algorithm - GA) hoặc các phương pháp tối ưu hóa khác. Hệ thống cần có khả năng:

- Tự động lập lịch học hợp lý: Phân bổ môn học, giảng viên và phòng học sao cho không vi phạm ràng buộc cứng, đồng thời tối ưu hóa theo các ràng buộc mềm.
- Giảm thiểu xung đột: Đảm bảo không có xung đột về giảng viên, phòng học và lớp học.
- Đảm bảo tính linh hoạt: Hệ thống có thể dễ dàng điều chỉnh khi có sự thay đổi về số lượng lớp học, giảng viên, phòng học hoặc môn học.
- Tăng hiệu suất sử dụng tài nguyên: Tận dụng tối đa các phòng học và giảng viên một cách hiệu quả.
- Tối ưu trải nghiệm của giảng viên và sinh viên: Phân bổ lịch học sao cho phù hợp với yêu cầu giảng dạy của giáo viên và đảm bảo sinh viên có lịch học hợp lý, tránh quá tải.
- Tự động cập nhật và thích ứng: Khi có thay đổi dữ liệu đầu vào (ví dụ: giảng viên mới, lớp học mới, phòng học thay đổi), hệ thống có thể nhanh chóng tạo ra một lịch học tối ưu mới mà không cần sắp xếp lại từ đầu.

Hệ thống cần có giao diện thân thiện để người dùng có thể xem, chỉnh sửa lịch học nếu cần, đồng thời cung cấp các công cụ đánh giá chất lượng lịch học dựa trên các tiêu chí tối ưu đã đề ra.

2.2 Mô tả thuật toán sử dụng

2.2.1 Giới thiệu thuật toán di truyền (Genetic Algorithm - GA)

Thuật toán di truyền (Genetic Algorithm - GA) là một phương pháp tối ưu hóa heuristic dựa trên nguyên lý chọn lọc tự nhiên trong tiến hóa sinh học. Thuật toán mô phỏng quá trình tiến hóa để tìm ra lời giải tốt nhất cho một bài toán cụ thể. Các thành phần chính của GA bao gồm:

- **Cá thể (Individual):** Mỗi cá thể đại diện cho một lời giải tiềm năng của bài toán.
- **Quần thể (Population):** Tập hợp nhiều cá thể, mỗi cá thể là một lời giải khả dĩ.
- **Hàm thích nghi (Fitness Function):** Xác định chất lượng của từng cá thể dựa trên mức độ phù hợp với mục tiêu cần tối ưu hóa.
- **Lựa chọn (Selection):** Chọn ra những cá thể có giá trị thích nghi cao nhất để tham gia vào quá trình sinh sản.
- **Lai ghép (Crossover):** Kết hợp thông tin từ hai cá thể cha mẹ để tạo ra cá thể con, giúp đa dạng hóa quần thể.
- **Đột biến (Mutation):** Thay đổi ngẫu nhiên một số gene của cá thể để duy trì tính đa dạng và tránh rơi vào cực tiểu cục bộ.
- **Tiến hóa qua nhiều thế hệ:** Thực hiện liên tục quá trình trên để tạo ra những cá thể có chất lượng ngày càng cao, tiến gần đến lời giải tối ưu.

2.2.2 Áp dụng thuật toán di truyền vào bài toán lập lịch

Trong bài toán lập lịch, mục tiêu là tìm ra một lịch học tối ưu sao cho các lớp học được phân bổ hợp lý theo các tiêu chí về thời gian, giảng viên, phòng học và các ràng buộc khác. Thuật toán di truyền được áp dụng theo các bước sau:

1. **Mã hóa nhiễm sắc thể (Chromosome Encoding):** Mỗi cá thể trong quần thể biểu diễn một lịch học khả dĩ. Một cá thể được mã hóa dưới dạng danh sách các gene, trong đó mỗi gene tương ứng với một lớp học, bao gồm các thông tin như:
 - Môn học (Subject)
 - Giảng viên (Lecturer)

- Thời gian học (Time Slot)
- Phòng học (Classroom)

Mỗi gene có thể được biểu diễn dưới dạng một bộ dữ liệu (subject_id, lecturer_id, time_slot, room_id), đảm bảo rằng toàn bộ lịch học được định nghĩa đầy đủ.

2. **Hàm thích nghi (Fitness Function):** Hàm thích nghi đánh giá mức độ phù hợp của một lịch học dựa trên số lượng ràng buộc bị vi phạm. Một lịch học tốt là lịch có ít xung đột nhất. Các tiêu chí để đánh giá có thể bao gồm:

- Không có trùng lặp phòng học: Một phòng không thể có nhiều lớp học cùng một thời điểm.
- Không có trùng lặp giảng viên: Một giảng viên không thể dạy nhiều lớp cùng lúc.
- Không có trùng lịch sinh viên: Sinh viên không thể học hai môn trong cùng một thời điểm.
- Ưu tiên thời gian mong muốn của giảng viên và sinh viên: Giảng viên có thể có thời gian giảng dạy ưu tiên, và sinh viên có thời gian rảnh nhất định.
- Cân bằng tải giảng viên: Giảng viên không nên có số tiết dạy quá chênh lệch nhau.

Điểm thích nghi của một cá thể sẽ được tính dựa trên tổng số ràng buộc bị vi phạm, với mục tiêu giảm thiểu số vi phạm để tối ưu lịch học.

3. **Toán tử lựa chọn (Selection Operator):** Các cá thể có điểm thích nghi cao sẽ được ưu tiên chọn để tạo ra thế hệ mới. Một số phương pháp chọn lọc phổ biến gồm:

- Chọn lọc tỷ lệ (Roulette Wheel Selection): Xác suất chọn một cá thể dựa trên giá trị thích nghi của nó.
- Chọn lọc giải đấu (Tournament Selection): Chọn ra một số cá thể ngẫu nhiên, sau đó lấy cá thể tốt nhất trong nhóm để lai ghép.
- Chọn lọc xếp hạng (Rank Selection): Sắp xếp các cá thể theo điểm thích nghi và chọn dựa trên thứ hạng.

4. **Toán tử lai ghép (Crossover Operator):** Sau khi chọn lọc, các cá thể cha mẹ được lai ghép để tạo ra thế hệ con. Một trong những phương pháp phổ biến là lai ghép một điểm (Single-Point Crossover):

- Chọn ngẫu nhiên một điểm cắt trong nhiễm sắc thể.
- Hoán đổi phần sau điểm cắt giữa hai cá thể cha mẹ để tạo ra hai cá thể con mới.

Ví dụ:

- Cha mẹ 1: (S1, L1, T1, R1) - (S2, L2, T2, R2) - (S3, L3, T3, R3)
- Cha mẹ 2: (S1, L1, T2, R3) - (S2, L2, T1, R1) - (S3, L3, T3, R2)

Sau lai ghép:

- Con 1: (S1, L1, T1, R1) - (S2, L2, T1, R1) - (S3, L3, T3, R2)
- Con 2: (S1, L1, T2, R3) - (S2, L2, T2, R2) - (S3, L3, T3, R3)

5. **Toán tử đột biến (Mutation Operator):** Mục đích của đột biến là tăng tính đa dạng của quần thể và tránh bị rơi vào cực tiểu cục bộ. Một số chiến lược đột biến bao gồm:

- Hoán đổi thời gian học: Thay đổi ngẫu nhiên slot thời gian của một lớp học.
- Hoán đổi phòng học: Đổi phòng của hai lớp trong cùng một khoảng thời gian.
- Hoán đổi giảng viên: Chọn ngẫu nhiên một giảng viên khác có thể giảng dạy môn học đó.

Ví dụ:

- Trước đột biến: (S1, L1, T1, R1) - (S2, L2, T2, R2)
- Sau đột biến (đổi phòng học): (S1, L1, T1, R2) - (S2, L2, T2, R1)

6. **Tiến hóa qua nhiều thế hệ:** Sau khi thực hiện các bước trên, quần thể mới được tạo ra và quá trình tiếp tục lặp lại trong nhiều thế hệ cho đến khi đạt được tiêu chí dừng (ví dụ: số thế hệ tối đa hoặc không cải thiện sau một số lần lặp).

2.2.3 Kết luận

Thuật toán di truyền là một phương pháp mạnh mẽ để giải quyết bài toán lập lịch, giúp tìm ra lịch học tối ưu bằng cách mô phỏng quá trình tiến hóa tự nhiên. Việc sử dụng các toán tử lai ghép, đột biến và chọn lọc hợp lý giúp thuật toán khám phá nhiều phương án khác nhau và tối ưu hóa hiệu quả lịch học dựa trên các ràng buộc đặt ra.

2.3 Phân tích mã nguồn thuật toán tạo lịch học tự động bằng thuật toán di truyền

2.3.1 Mô tả tổng quan về chương trình

Chương trình này sử dụng thuật toán di truyền để tự động tạo lịch học cho các lớp học, nhằm tối ưu hóa việc sắp xếp thời gian, phòng học và giáo viên. Thuật toán di truyền hoạt động bằng cách tạo ra nhiều lịch học ngẫu nhiên ban đầu, sau đó cải thiện chúng qua nhiều thế hệ bằng các bước chọn lọc, lai ghép và đột biến để tìm ra lịch học phù hợp nhất.

Mục tiêu chính:

- Phân bổ thời gian học hợp lý, tránh chồng chéo lịch.
- Đảm bảo không có xung đột lịch giữa giáo viên và sinh viên.
- Bố trí phòng học phù hợp với sức chứa và yêu cầu về thiết bị.
- Tối ưu hoá sự phân bổ thời gian giảng dạy theo mong muốn của giáo viên.

2.3.2 Các lớp đối tượng chính trong chương trình

Lớp Class – Đại diện cho lớp học

Lớp này chứa thông tin về môn học, giáo viên giảng dạy, danh sách sinh viên, thời gian học và các thiết bị cần thiết. Các thuộc tính chính:

- Môn học.
- Giáo viên giảng dạy.
- Danh sách sinh viên.
- Thời lượng buổi học.
- Danh sách thiết bị cần thiết.

Lớp Room – Đại diện cho phòng học

Lớp này chứa thông tin về phòng học, bao gồm sức chứa, thiết bị có sẵn và danh sách các khung giờ trống. Các thuộc tính chính:

- Sức chứa tối đa.
- Danh sách thiết bị có sẵn.

- Danh sách các khung giờ trống.

Lớp Teacher – Đại diện cho giáo viên

Lớp này chứa thông tin về giáo viên, bao gồm danh sách các môn học có thể giảng dạy và thời gian có thể dạy. Các thuộc tính chính:

- Tên giáo viên.
- Danh sách môn học có thể giảng dạy.
- Danh sách thời gian có thể giảng dạy.

Lớp Schedule – Đại diện cho một lịch học

Lớp này quản lý toàn bộ thời khóa biểu, bao gồm danh sách các lớp học, phòng học và thời gian học. Các thuộc tính và phương thức quan trọng:

- Danh sách các lớp đã được xếp lịch.
- Phương thức thêm một lớp vào lịch học.
- Phương thức đánh giá mức độ phù hợp của lịch học dựa trên các tiêu chí:
 - Xung đột lịch giữa giáo viên.
 - Xung đột lịch giữa sinh viên.
 - Phòng học có đủ sức chứa không.
 - Phòng học có đủ thiết bị cần thiết không.
 - Mức độ phù hợp với thời gian mong muốn của giáo viên.

2.3.3 Cách hoạt động của thuật toán di truyền trong chương trình

Khởi tạo quần thể

Ban đầu, chương trình tạo một tập hợp các lịch học ngẫu nhiên để đảm bảo sự đa dạng trong việc tìm kiếm giải pháp tốt nhất.

Đánh giá hàm thích nghi (Fitness Function)

Mỗi lịch học được đánh giá dựa trên số lượng ràng buộc bị vi phạm. Lịch học có số lỗi càng ít sẽ có điểm thích nghi càng cao. Các tiêu chí đánh giá:

- Tránh xung đột lịch giữa giáo viên.

- Tránh xung đột lịch giữa sinh viên.
- Đảm bảo phòng học có đủ sức chứa và thiết bị cần thiết.
- Đảm bảo thời gian giảng dạy phù hợp với mong muốn của giáo viên.

Chọn lọc cá thể tốt nhất (Selection)

Dựa trên điểm thích nghi, các lịch học tốt nhất sẽ được chọn để tiếp tục tạo ra thế hệ sau. Các phương pháp chọn lọc:

- Tournament Selection: Chọn nhóm ngẫu nhiên các cá thể và giữ lại cá thể tốt nhất.
- Roulette Wheel Selection: Chọn cá thể dựa trên xác suất tỉ lệ thuận với điểm thích nghi.

Lai ghép (Crossover) để tạo lịch mới

Hai lịch học tốt nhất được lai ghép để tạo ra lịch học mới bằng cách kết hợp thông tin từ cả hai lịch cha mẹ. Nguyên tắc lai ghép:

- Chọn một số lớp từ lịch cha và một số từ lịch mẹ.
- Đảm bảo lịch con không vi phạm các ràng buộc.

Đột biến (Mutation) để tăng đa dạng

Một số lịch học sẽ bị thay đổi ngẫu nhiên để tăng khả năng tìm ra giải pháp tốt hơn. Các loại đột biến có thể xảy ra:

- Thay đổi thời gian học của một lớp.
- Chuyển lớp sang phòng học khác.
- Thay đổi giáo viên giảng dạy.

Kiểm tra điều kiện dừng

Thuật toán tiếp tục qua nhiều thế hệ cho đến khi đạt một trong các điều kiện dừng:

- Lịch học đạt điểm thích nghi đủ cao.
- Số thế hệ đạt giới hạn tối đa.

2.3.4 Khởi tạo dữ liệu đầu vào

Dữ liệu đầu vào bao gồm danh sách giáo viên, lớp học và phòng học:

- **Danh sách giáo viên:**

- Giáo viên với danh sách môn học có thể giảng dạy.
- Danh sách thời gian mà giáo viên có thể dạy.

- **Danh sách lớp học:**

- Các lớp học với môn học, giáo viên, danh sách sinh viên.
- Yêu cầu thiết bị của lớp học.

- **Danh sách phòng học:**

- Phòng học với sức chứa và danh sách thiết bị có sẵn.
- Các khung giờ mà phòng học còn trống.

2.3.5 Chạy thuật toán và hiển thị kết quả

Sau khi thuật toán di truyền kết thúc, chương trình hiển thị lịch học tối ưu nhất. Thông tin hiển thị trong lịch học:

- Môn học.
- Giáo viên giảng dạy.
- Phòng học được sử dụng.
- Thời gian học.

Nhờ vào thuật toán di truyền, chương trình có thể tự động sắp xếp lịch học một cách hiệu quả, tiết kiệm thời gian và tối ưu hoá việc sử dụng tài nguyên giảng dạy.

Chương 3

Kết luận

3.1 Đánh giá chung về đề tài

Đề tài “Xây dựng chương trình tạo lịch học tự động bằng thuật toán di truyền” là một ứng dụng thực tiễn của trí tuệ nhân tạo vào lĩnh vực giáo dục. Việc lập lịch học tự động giúp giảm tải công việc thủ công, tiết kiệm thời gian và tối ưu hóa việc sử dụng tài nguyên như giáo viên, phòng học và khung giờ học.

Sử dụng thuật toán di truyền (Genetic Algorithm - GA) giúp tìm kiếm lịch học hợp lý trong một không gian tìm kiếm lớn, đảm bảo tính linh hoạt và khả năng thích nghi với các yêu cầu khác nhau. Mặc dù đây chưa phải là phương pháp duy nhất để giải quyết bài toán lập lịch, nhưng thuật toán di truyền mang lại lợi ích lớn trong việc xử lý các ràng buộc phức tạp và tối ưu hóa nhiều tiêu chí cùng lúc.

3.1.1 Những kết quả đạt được

- **Tự động hóa quá trình lập lịch**
 - Giảm tải công việc thủ công, hạn chế sai sót do con người.
 - Đảm bảo lịch học hợp lý theo các ràng buộc có sẵn.
- **Giải quyết xung đột lịch**
 - Hạn chế tình trạng trùng lịch giữa giáo viên, sinh viên và phòng học.
 - Phân bổ phòng học phù hợp với số lượng sinh viên và thiết bị yêu cầu.
- **Tối ưu hóa tài nguyên**

- Tận dụng tối đa khung giờ trống của giáo viên và phòng học.
- Cân đối số lượng lớp học trong từng khung giờ.
- **Tính linh hoạt cao**
 - Có thể dễ dàng điều chỉnh dữ liệu đầu vào để phù hợp với từng trường học.
 - Thích ứng với số lượng lớp học, giáo viên và phòng học khác nhau.

3.1.2 Một số hạn chế của đề tài

- **Không đảm bảo tìm được lời giải tối ưu tuyệt đối**
 - Thuật toán di truyền chỉ tìm kiếm lời giải gần tối ưu, có thể không phải là phương án tốt nhất trong mọi trường hợp.
- **Thời gian xử lý có thể cao khi số lượng lớp học lớn**
 - Khi số lượng lớp học, giáo viên và phòng học tăng lên, thuật toán phải kiểm tra nhiều khả năng kết hợp hơn, dẫn đến thời gian tính toán lâu hơn.
- **Chưa tối ưu hóa các thông số của thuật toán di truyền**
 - Các tham số như tỷ lệ lai ghép (crossover rate), tỷ lệ đột biến (mutation rate), kích thước quần thể (population size) có thể chưa được điều chỉnh tối ưu, ảnh hưởng đến tốc độ hội tụ của thuật toán.
- **Chưa tích hợp giao diện người dùng trực quan**
 - Hiện tại, chương trình chỉ hiển thị kết quả dưới dạng văn bản, chưa có giao diện đồ họa để người dùng dễ dàng tương tác.

3.2 Hướng phát triển của đề tài

- **Cải tiến thuật toán để nâng cao hiệu suất**
 - Kết hợp với các thuật toán tối ưu khác như thuật toán bầy đàn (PSO) hoặc thuật toán mô phỏng tôi luyện (Simulated Annealing - SA) để cải thiện tốc độ hội tụ.
 - Tối ưu tham số thuật toán bằng cách sử dụng học tăng cường (Reinforcement Learning) để tự động điều chỉnh các tham số của thuật toán di truyền.
- **Tích hợp thêm các ràng buộc thực tế**

- Hỗ trợ các ưu tiên của giáo viên, như mong muốn giảng dạy vào buổi sáng hay chiều.
- Xử lý trường hợp đặc biệt như giáo viên dạy ở nhiều cơ sở hoặc lớp học cần thiết bị chuyên biệt.
- **Xây dựng giao diện người dùng trực quan**
 - Phát triển ứng dụng web hoặc phần mềm desktop để người dùng dễ dàng nhập dữ liệu, theo dõi quá trình chạy thuật toán và xem kết quả trực quan hơn.
- **Mở rộng phạm vi ứng dụng**
 - Áp dụng thuật toán vào lập lịch thi, lập lịch giảng dạy cho trường đại học hoặc lập lịch làm việc cho nhân viên trong doanh nghiệp.

3.3 Kết luận

Đề tài đã đạt được những kết quả đáng kể trong việc ứng dụng thuật toán di truyền vào bài toán lập lịch học, giúp tự động hóa, giải quyết xung đột, và tối ưu hóa tài nguyên. Tuy nhiên, vẫn còn một số hạn chế cần khắc phục, đặc biệt là về tối ưu hiệu suất thuật toán và phát triển giao diện trực quan. Với các hướng cải tiến đã đề xuất, chương trình có thể được hoàn thiện hơn, nâng cao hiệu quả và mở rộng phạm vi ứng dụng trong tương lai.