**ỦY BAN NHÂN DÂN TP HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC SÀI GÒN**

**KHOA CÔNG NGHỆ THÔNG TIN**

****

**Họ và tên sinh viên: Lê Minh Cường**

**MSSV : 3117410027**

**ĐỀ XUẤT THUẬT TOÁN HILL CLIMBING GIẢI BÀI TOÁN NGƯỜI BÁN HÀNG**

**Giảng viên hướng dẫn: Phan Tấn Quốc**

***TP. Hồ Chí Minh, tháng ….. năm 2021***

**MỤC LỤC**

[CHƯƠNG 1. TỔNG QUAN VỀ BÀI TOÁN NGƯỜI BÁN HÀNG 5](#_Toc69652694)

[1.1. Một số định nghĩa 5](#_Toc69652695)

[1.1.1. Bài toán người bán hàng 5](#_Toc69652696)

[1.1.2. Định nghĩa 2: Bài toán người bán hàng trong đồ thị 5](#_Toc69652697)

[1.1.3. Hệ đo lường 5](#_Toc69652698)

[1.1.4. Euclidean 6](#_Toc69652699)

[1.1.6. Độ phức tạp của thuật toán 6](#_Toc69652700)

[1.1.7. Các trường hợp đặc biệt 7](#_Toc69652701)

[1.1.7.1. Hệ đo lường và tổng quan: 7](#_Toc69652702)

[1.1.7.2. Đi qua lặp lại và không lặp lại 7](#_Toc69652703)

[1.1.8. Ví dụ về một trường hợp với đồ thị đối xứng 7](#_Toc69652704)

[1.2. Ứng dụng của bài toán người bán hàng 7](#_Toc69652705)

[1.3. Lịch sử nghiên cứu vấn đề/ Tổng quan 8](#_Toc69652706)

[1.4. Kết luận 9](#_Toc69652707)

[2. CHƯƠNG 2 NGHIÊN CỨU VỀ THUẬT TOÁN HILL CLIMBING 10](#_Toc69652708)

[2.1. Định nghĩa chung về Hill Climbing 10](#_Toc69652709)

[2.2. Mô tả toán học 10](#_Toc69652710)

[2.3. Đặc điểm của bài toán Hill Climbing 11](#_Toc69652711)

[2.4. Đặc điểm các vùng trong Hill Climbing 11](#_Toc69652712)

[2.4.1. Local maximum 11](#_Toc69652713)

[2.4.2. Global maximum 12](#_Toc69652714)

[2.4.3. Plateua/ flat local maximum 12](#_Toc69652715)

[2.4.4. Shoulder 12](#_Toc69652716)

[2.5. Các loại Hill Climbing 13](#_Toc69652717)

[2.5.1. Simple Hill Climbing 13](#_Toc69652718)

[2.5.2. Steepest Ascent Hill Climbing 13](#_Toc69652719)

[2.5.3. Stochastic Hill Climbing 13](#_Toc69652720)

[2.6. Kết luận chương 2 13](#_Toc69652721)

[3. CHƯƠNG 3: ĐỀ XUẤT THUẬT TOÁN HILL CLIMBING GIẢI BÀI TOÁN NGƯỜI BÁN HÀNG 15](#_Toc69652722)

[3.1. Cơ sở lý thuyết 15](#_Toc69652723)

[3.2. Giải bài toán người bán hàng 15](#_Toc69652724)

[3.2.1. Trình bài giải thuật 15](#_Toc69652725)

[3.2.1.1. Khởi tạo lời giải ban đầu 15](#_Toc69652732)

[3.2.1.2. Xác định điều kiện dừng của bài toán 18](#_Toc69652733)

[3.2.1.3. Khởi tạo tuyến đường S’ và đánh giá với tuyến đường S 18](#_Toc69652734)

[3.2.1.4. Tổng hợp tất cả các bước 21](#_Toc69652735)

[3.2.2. Giải bài toán với giải thuật Hill Climbing 21](#_Toc69652736)

[4. CHƯƠNG 4 THỰC NGHIỆM ĐÁNH GIÁ 25](#_Toc69652737)

[4.1. Môi trường thực nghiệm 25](#_Toc69652738)

[4.1.1. Về máy tính: 25](#_Toc69652739)

[4.1.1. Về ngôn ngữ: 25](#_Toc69652740)

[4.2. Hệ thống thực nghiệm 25](#_Toc69652741)

[4.3. Kết quả thực nghiệm 25](#_Toc69652742)

[4.4. Đánh giá chất lượng lời giải và kết quả 26](#_Toc69652743)

[4.5. Kết luận và đóng góp. 26](#_Toc69652744)

[TÀI LIỆU THAM KHẢO 27](#_Toc69652745)

**XÁC NHẬN CỦA NGƯỜI HƯỚNG DẪN KHOA HỌC**

**……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………….……………………………………………………………………………………….……………………………………………………………………………………….……………………………………………………………………**

|  |  |
| --- | --- |
|  | *Thành phố Hồ Chí Minh, ngày tháng năm*  **Người hướng dẫn khoa học**  *(Kí và ghi rõ họ tên)*  **TS. Phan Tấn Quốc** |

# 

# CHƯƠNG 1. TỔNG QUAN VỀ BÀI TOÁN NGƯỜI BÁN HÀNG

## Một số định nghĩa

### Bài toán người bán hàng

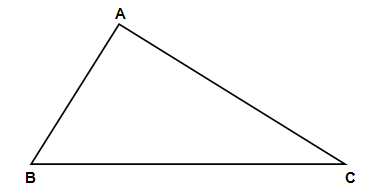
Bài toán người bán hàng (còn được biết tới với tên là nhân viên bán hàng hay viết tắt là TSP) có nội dung cụ thể như sau: Cho danh sách tất cả các thành phố và khoảng cách của mỗi thành phố, bài toán muốn tìm ra tuyến đường đi ngắn nhất có thể sao cho đi qua mỗi thành phố một lần và quay trở về thành phố ban đầu sao cho chi phí là nhỏ nhất. Đây là bài toán thuộc lớp NP-hard trong tối ưu hóa tổ hợp, quan trọng trong lý thuyết khoa học máy tính và hoạt động nghiên cứu.

### Định nghĩa 2: Bài toán người bán hàng trong đồ thị

Bài toán người bán hàng có thể được mô hình hóa như là một đồ thị vô hướng có trọng số, vì vậy những thành phố sẽ trở thành đỉnh, những đường đi sẽ trở thành cạnh của đồ thị và khoảng cách của đường đi sẽ trở thành trọng số. Vấn đề trở thành việc bắt đầu và kết thúc tại 1 đỉnh được chỉ đỉnh sau khi đi qua tất cả các đỉnh chỉ một lần. Đây là mô hình đồ thị hoàn chỉnh. Nếu không tồn tại đường đi giữa 2 thành phố, việc thêm đủ cạnh vào để hoàn thành đồ thị sẽ không ảnh hưởng đến kết quả của bài toán.

### Hệ đo lường

Trong hệ đo lường TSP, còn được biết tới như delta-TSP hay là , khoảng cách thỏa bất đẳng thức hình tam giác, có nghĩa là khoảng cách từ A đến B sẽ không bao giờ dài hơn khoảng cách thông qua C

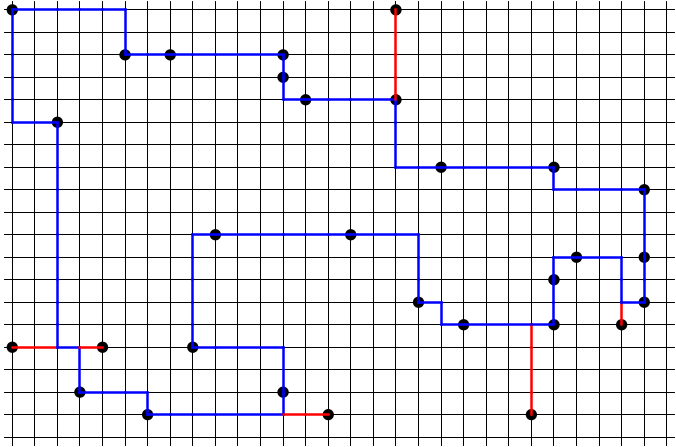


Bài toán TSP trong các hệ đo lường khác nhau. Ví dụ:

* Trong hệ đo lường Euclidean khoảng cách từ 2 thành phố là khoảng cách Euclidean giữa 2 điểm tương ứng

Công thức Euclidean giữa 2 điểm bất kì là d = với là chiều của x và chiều y của 2 điểm.

* Trong  rectilinear TSP thì khoảng cách giữa 2 thành phố là tổng trị tuyệt đối giá trị của sự khác nhau của tọa độ x và y



Công thức khoảng cách Mahattan giữa 2 điểm bất kì là d =  với là chiều của x và chiều y của 2 điểm.

### Euclidean

Đầu vào có thể là những con số là số thực tùy ý, Euclidean TSP là một trường hợp điển hình của bài toán trong hệ đo lường TSP, bởi vì khoảng cách trong không gian xy tuân theo bất đẳng thức tam giác.

* + 1. **Định nghĩa 3: Đối xứng và bất đối xứng**

Trong bài toán người bán hàng với đồ thị đối xứng thì khoảng cách giữa 2 thành phố là như nhau với mỗi chiều ngược lại, ví dụ như: đồ thị vô hướng. Bài toán với đồ thị đối xứng thì có thể có số lượng giải pháp có thể. Trong khi đồ thị bất đối xứng, đường đi có thể không tồn tại cả 2 chiều và khoảng cách cũng có thể khác nhau, ví dụ: đồ thị có hướng. Hầu hết tất cả khoảng cách giữa hai thành phố trong mạng lưới TSP thì như nhau ở cả 2 chiều tức là A đến B bằng khoảng cách từ B đến A. Trường hợp khi mà khoảng cách từ A đến B không bằng khoảng cách từ B đến A đó được gọi là không đối xứng.

### Độ phức tạp của thuật toán

Bài toán có độ phức tạp là NP-hard không thể giải bài toán trực tiếp với dữ liệu đầu vào lớn một cách chính xác, mà chỉ có thể giải bài toán gần đúng. Vấn đề sẽ luôn luôn là NP-hard trong trường hợp khi mà những thành phố nằm trong trục tọa độ xy với hệ khoảng cách tính bằng Eclidean.

### Các trường hợp đặc biệt

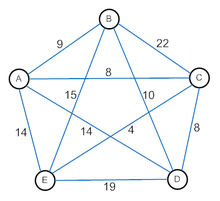
#### Hệ đo lường và tổng quan:

Trong phiên bản đo lường của TSP, công thức khoảng cách d là một hệ đo lường. Nó thỏa với bất đẳng thức tam giác tức là nếu chúng ta có 3 điểm A, B và C muốn đi từ A qua C chúng ta có thể đi một đường khác là A sang B rồi từ B sang C. Như vậy thì sẽ luôn luôn thỏa bất đẳng thức d(AC) <= d(AB) + d(BC). Trong những phiên bản khác, d có thể được gán bất kì giá trị nào của một cạnh, nghĩa là từ A qua B có thể không cần từ B qua A không tuân theo bất đẳng thức tam giác.

#### Đi qua lặp lại và không lặp lại

Mục tiêu của TSP là tìm một chu trình đi qua tất cả các cạnh chỉ một lần. Vậy thì có chu trình nào chứa lặp lại đỉnh mà đã đi qua được hay không? Trong phiên bản TSP không lặp lại thì, chu trình phải đi qua mỗi thành phố chỉ một lần duy nhất. Nhưng trong phiên bản lặp lại thì chúng ta có thể chấp nhận được việc lặp lại việc đi qua lại thành phố thêm một lần nữa, nếu kết quả là tốt hơn.

### Ví dụ về một trường hợp với đồ thị đối xứng



Hình 1. Minh họa đồ thị và đường đi bắt đầu từ đỉnh C {C, E, A, B, D, C}

Cho đồ thị vô hướng liên thông có 9 đỉnh và 26 cạnh như Hình 1; Lấy một đường đi bất kì tổng chi phí của đoạn đường bắt đầu từ C gồm CEABDC là 4 + 14 + 9 + 10 + 8 = 45. Vậy kết quả cuối cùng đi từ C là 45 nếu chúng ta ưu tiên đi những đường ngắn từ đỉnh hiện tại tới đỉnh tiếp theo.

## Ứng dụng của bài toán người bán hàng

Ngoài việc là một "polytope" của một vấn đề tối ưu hóa tổ hợp khó khăn từ một phức tạp điểm lý thuyết của xem, có những trường hợp quan trọng của các vấn đề thực tế có thể được xây dựng như các vấn đề TSP và nhiều vấn đề khác là những khái quát của vấn đề này.

Bên cạnh việc khoan mạch in bảng mô tả ở trên, vấn đề có cấu trúc TSP xảy ra trong phân tích cấu trúc của các tinh thể, (Bland và Shallcross, 1987), các đại tu động cơ tuốc bin khí (Pante, Lowe và Chandrasekaran, 1987), trong xử lý vật liệu trong một nhà kho (Ratliff và Rosenthal, 1981) , trong việc cắt giảm các vấn đề chứng khoán, (Garfinkel, 1977), các phân nhóm của các mảng dữ liệu, (Lenstra và Rinooy Kạn, 1975), trình tự các công việc trên một máy tính duy nhất (và Gilmore Gomory, 1964) và phân công các tuyến đường cho máy bay của một hạm đội quy định (Boland, Jones, và Nemhauser, 1994).

Biến thể có liên quan về vấn đề nhân viên bán hàng đi du lịch bao gồm các nguồn tài nguyên hạn chế đi du lịch vấn đề nhân viên bán hàng trong đó có các ứng dụng trong lập kế hoạch với thời hạn tổng hợp (Pekny và Miller, 1990). Nghiên cứu này cũng cho thấy giải thưởng thu thập đi vấn đề nhân viên bán hàng (Balas, 1989) và các vấn đề Orienteering (Golden, Levy và Vohra, 1987) là trường hợp đặc biệt của tài nguyên hạn chế TSP.

Quan trọng nhất là vấn đề nhân viên bán hàng đi du lịch thường thể hiện như một bài toán con trong nhiều vấn đề tổ hợp phức tạp, là nổi tiếng và quan trọng nhất trong số đó là vấn đề định tuyến xe, có nghĩa là, vấn đề xác định cho một đội xe mà khách hàng sẽ được phục vụ bởi mỗi chiếc xe và theo thứ tự mỗi chiếc xe nên đến các khách hàng được giao. Đối với các cuộc điều tra có liên quan, xem Christofides (1985) và Fisher (1987).

## Lịch sử nghiên cứu vấn đề/ Tổng quan

Nguồn gốc của bài toán người bán hàng vẫn chưa được biết rõ. Một cuốn sổ tay dành cho người bán hàng xuất bản năm 1832 có đề cập đến bài toán này và có ví dụ cho chu trình trong nước Đức và Thụy Sĩ, nhưng không chứa bất kì nội dung toán học nào.

Bài toán người bán hàng được định nghĩa trong thế kỉ 19 bởi nhà toán học Ireland [William Rowan Hamilton](https://vi.wikipedia.org/w/index.php?title=W._R._Hamilton&action=edit&redlink=1) và nhà toán học Anh [Thomas Kirkman](https://vi.wikipedia.org/w/index.php?title=Thomas_Kirkman&action=edit&redlink=1). [Trò chơi Icosa](https://vi.wikipedia.org/w/index.php?title=Tr%C3%B2_ch%C6%A1i_Icosa&action=edit&redlink=1) của Hamilton là một trò chơi giải trí dựa trên việc tìm kiếm [chu trình Hamilton](https://vi.wikipedia.org/wiki/%C4%90%C6%B0%E1%BB%9Dng_%C4%91i_Hamilton). Trường hợp tổng quát của TSP có thể được nghiên cứu lần đầu tiên bởi các nhà toán học ở Vienna và Harvard trong những năm 1930, đặc biệt là [Karl Menger](https://vi.wikipedia.org/w/index.php?title=Karl_Menger&action=edit&redlink=1), người đã định nghĩa bài toán, xem xét thuật toán hiển nhiên nhất cho bài toán, và phát hiện ra thuật toán láng giềng gần nhất là không tối ưu.

[Hassler Whitney](https://vi.wikipedia.org/w/index.php?title=Hassler_Whitney&action=edit&redlink=1) ở [đại học Princeton](https://vi.wikipedia.org/wiki/%C4%90%E1%BA%A1i_h%E1%BB%8Dc_Princeton) đưa ra tên bài toán người bán hàng ngay sau đó.

Trong những năm 1950 và 1960, bài toán trở nên phổ biến trong giới nghiên cứu khoa học ở châu Âu và Mỹ. [George Dantzig](https://vi.wikipedia.org/wiki/George_Dantzig), [Delbert Ray Fulkerson](https://vi.wikipedia.org/wiki/Delbert_Ray_Fulkerson) và Selmer M. Johnson ở công ty RAND tại Santa Monica đã có đóng góp quan trọng cho bài toán này, biểu diễn bài toán dưới dạng [quy hoạch nguyên](https://vi.wikipedia.org/w/index.php?title=Quy_ho%E1%BA%A1ch_nguy%C3%AAn&action=edit&redlink=1) và đưa ra phương pháp [mặt phẳng cắt](https://vi.wikipedia.org/w/index.php?title=M%E1%BA%B7t_ph%E1%BA%B3ng_c%E1%BA%AFt&action=edit&redlink=1) để tìm ra lời giải. Với phương pháp mới này, họ đã giải được tối ưu một trường hợp có 49 thành phố bằng cách xây dựng một chu trình và chứng minh rằng không có chu trình nào ngắn hơn. Trong những thập niên tiếp theo, bài toán được nghiên cứu bởi nhiều nhà nghiên cứu trong các lĩnh vực toán học, khoa học máy tính, hóa học, vật lý, và các ngành khác.

Năm 1972, [Richard M. Karp](https://vi.wikipedia.org/w/index.php?title=Richard_M._Karp&action=edit&redlink=1) chứng minh rằng bài toán [chu trình Hamilton](https://vi.wikipedia.org/wiki/%C4%90%C6%B0%E1%BB%9Dng_%C4%91i_Hamilton) là [NP-đầy đủ](https://vi.wikipedia.org/wiki/NP-%C4%91%E1%BA%A7y_%C4%91%E1%BB%A7), kéo theo bài toán TSP cũng là [NP-đầy đủ](https://vi.wikipedia.org/wiki/NP-%C4%91%E1%BA%A7y_%C4%91%E1%BB%A7). Đây là một lý giải toán học cho sự khó khăn trong việc tìm kiếm chu trình ngắn nhất.

Một bước tiến lớn được thực hiện cuối thập niên 1970 và 1980 khi Grötschel, Padberg, Rinaldi và cộng sự đã giải được những trường hợp lên tới 2392 thành phố, sử dụng phương pháp mặt phẳng cắt và [nhánh cận](https://vi.wikipedia.org/w/index.php?title=Nh%C3%A1nh_c%E1%BA%ADn&action=edit&redlink=1).

## Kết luận

Bài toán người bán hàng là một trong những vấn đề hay kể cả trong toán học và các lĩnh vực nghiên cứu khoa học, không chỉ góp phần vào việc xây dựng các đường đi ngắn nhất giữa các thành phố, làm tối ưu hóa cách giải nghĩa là cải thiện được tuyến đường đi trên thực tế. Với độ phức tạp là NP hard như vậy sẽ cần có những cách giải và lời giải hợp lý nếu không bài toán sẽ không thể nào giải ra được kết quả vì có rất nhiều trường hợp.

# CHƯƠNG 2 NGHIÊN CỨU VỀ THUẬT TOÁN HILL CLIMBING

## Định nghĩa chung về Hill Climbing

Hill Climbing là một kĩ thuật tối ưu hóa bài toán, thuộc cùng nhóm với bài toán local search. Hill Climbing là một heuristic search được dùng cho những vấn đề tối ưu hóa toán học trong lĩnh vực Trí Tuệ Nhân Tạo. Nó là thuật toán lặp (iterative algorithm), tức là là nó sẽ bắt đầu tại một lời giải bất kì tức nhiên lời giải đó không phải là một lời giải tối ưu (global maximums) và cho đó là một lời giải hiện tại của bài toán, sau đó nó sẽ cố gắng tìm một lời giải tốt hơn lời bằng việc tạo ra sự thay đổi trong lời giải hiện tại. Nếu nó tạo ra một lời giải tốt hơn, thì lời giải mới này sẽ được cho là lời giải hiện tại và sẽ lặp lại cho đến khi mà kết quả lời giải không thể tốt hơn được nữa.

* Trong định nghĩa trên, vấn đề tối ưu hóa toán học được tiến hành là hill climbing giải quyết vấn đề mà chỗ đó cần tối đa hóa hay giảm thiểu hóa một hàm cho trước bằng việc chọn một giá trị từ việc nhập giá trị đó vào hàm cho kết quả thỏa yêu cầu hay nhu cầu một phần của bài toán. Ví dụ ở đây là chúng ta muốn giảm thiểu hóa khoảng cách di chuyển của người bán hàng.
* Heuristic search có nghĩa là một thuật toán tìm kiếm không phải tìm kiếm tối ưu cách giải quyết vấn đề. Mà là nó sẽ tìm kiếm một cách giải quyết tốt trong một khoảng thời gian hợp lý.
* Hàm heuristic là hàm mà sẽ liệt kê và xếp hạng tất cả những phương án thay thế cho phương án hiện tại ở một bước, hay giai đoạn nào đó trong lúc tìm kiếm dựa trên những thông tin có sẵn, nó sẽ giúp thuật toán chọn ra được phương án tốt nhất trong những phương án thay thế đó. Giả dụ như tại bước A có thể đi qua B và C, thì thuật toán sẽ đánh giá là nên đi qua từ A qua B hay từ A qua C là tốt hơn.

## Mô tả toán học

Hill climbing cố gắng tối đa hóa (giảm thiểu hóa) hàm mục tiêu trong đó x là vector của biến liên tục hoặc rời rạc. Tại mỗi lần lặp, hill climbing sẽ chỉnh từng thuộc tính trong x và xác định bất cứ thay đổi nào cải thiện được giá trị của . Với hill climbing thì bất cứ sự thay đổi nào mà cải thiện được thì đều được chấp nhận, quá trình này sẽ được lặp cho tới khi mà không có sự thay đổi nào được tìm thấy để cải thiện được giá trị của . Khi đó thì x được gọi là locally optimal. Nếu hiểu theo nghĩa của bài toán TSP thì x là đường đi có thể đi và là tổng giá trị đường đi qua các đỉnh và về thành phố ban đầu. Và mục tiêu là cải thiện về giá trị nhỏ nhất

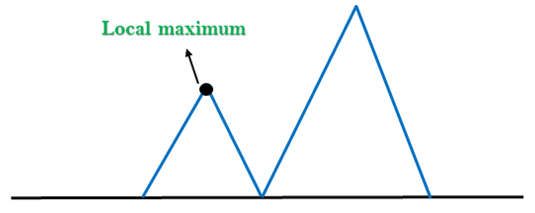
## Đặc điểm của bài toán Hill Climbing

* Nó tiến hành thuật toán tham lam: Có nghĩa là nó sẽ di chuyển về một hướng mà ở đó giá trị của hàm giá trị được tối ưu. Tiếp cận bằng giải thuật tham lam cho phép thuật toán tạo nên kết quả local maximum hoặc local minimum.
* Không đệ quy: Thuật Hill Climbing chỉ làm việc với giải pháp hiện tại (current solution) và giải phép có thể (impossible solution được tạo ra từ việc kế thừa kết quả của current solution). Nó không lấy lại kết quả của giải phá trước đó nên không cần quay lui để tìm lại.
* Feedback mechanism: Thuật toán có phản hồi kết kỹ thuật tức là nó giúp quyết định được hướng di chuyển (Khi nào thì nên leo lên và khi nào thì xuống núi). Những phản hồi này giúp được nâng cao thông qua việc tạo và thử nghiệm các chiến lược.
* Tăng dần sự thay đổi: Thuật toán giúp trao dồi giải pháp hiện tại bằng việc tang dần sự thay đổi. Có nghĩa là khi ở một trạng thái hiện tại, chúng ta làm nhiều bước để tìm ra tất cả những giải pháp có thể từ giải pháp hiện tại, sau đó chúng ta sẽ chọn giải pháp tốt nhất trong những giải pháp vừa mới sinh ra đó để so sánh với giải pháp hiện tại, nếu như tốt hơn thì sẽ thay thế giải pháp hiện tại. Cứ như vậy sẽ làm thay đổi nhiều bộ giải pháp theo giải pháp hiện tại.

## Đặc điểm các vùng trong Hill Climbing

### Local maximum

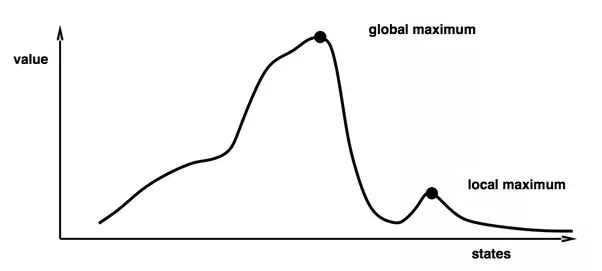
Là trạng thái tốt hơn tất cả những trạng thái lân cận. Tuy nhiên, vẫn tồn tại một trạng thái tốt nhất đó gọi là global maximum. Local maximum là tốt hơn bởi vì ở đây giá trị của hàm mục tiêu cao hơn những giá trị hàng xóm của nó.



Như hình có thể thấy rằng xung quanh trạng thái local maximum thì đó chính là trạng thái tốt nhất, nếu như mà thuật toán hill climbing không tốt thì khi tìm tất cả những trạng thái xung quanh local maximum thì chính là kết quả cuối cùng của bài toán, trong khi kết quả thực chất có thể sẽ tìm thấy đâu đó.

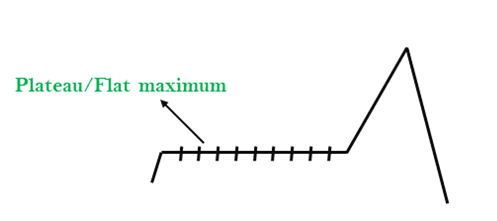
### Global maximum

Đây là trạng thái tốt nhất hay còn gọi là kết quả của bài toán. Ở trạng thái này thì đường đi, đi qua các thành phố là tốt nhất. Tức nhiên sẽ có nhiều trạng thái Global maximum tức nhiều đường đi.



### Plateua/ flat local maximum

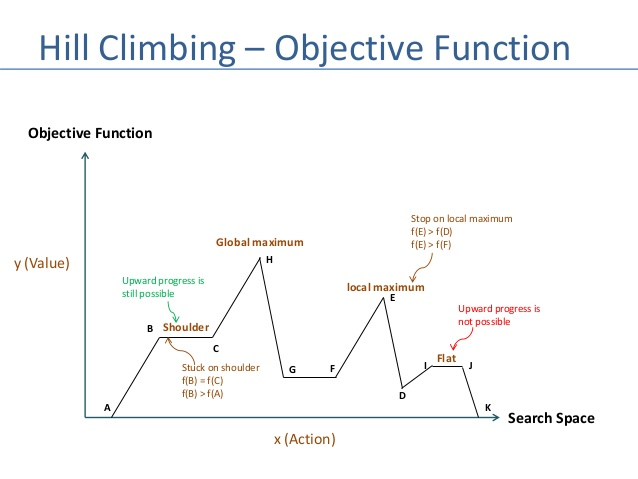
Là vùng mà của không gian trạng thái, mà ở đây tất cả các giá trị hàng xóm đề có kết quả là như nhau.



### Shoulder

Nó là một vùng cao nguyên có cạnh dốc nếu tiếp tục tìm trong vùng này vẫn có thể leo lên được dốc cao hơn để tìm ra được kết quả global maximum. Ở vùng này gần như đã gần chạm được tới được mục tiêu nếu như tiếp tục leo lên cao hơn.

Tổng kết lại chúng ta có được hình ảnh như dưới đây



## Các loại Hill Climbing

### Simple Hill Climbing

Nó sẽ tính toán từng lời giải lận cận theo thứ tự và chọn lời giải đầu tiên cái mà tối ưu lời giải hiện tại và lời giải tối ưu này sẽ được gán làm lời giải hiện tại. Và cứ thế tiếp tục lời giải tiếp theo thứ tự.

### Steepest Ascent Hill Climbing

Nó sẽ tính toán hết tất cả lời giải lân cận và chọn lời giải tốt nhất trong tất cả lời giải đó. Sau đó lời giải tốt đó sẽ được so sánh với lời giải hiện tại. Nếu tốt hơn sẽ được thay thế cho lời giải hiện tại. Cứ thế lặp lại liên tục cho đến khi trúng điều kiện dừng.

### Stochastic Hill Climbing

Nó sẽ không phải tính toán tất cả những lời giải lân cận. Thay vào đó nó sẽ tạo ngẫu nhiên một lời giải ngẫu nhiên và quyết định chọn lời giải đó làm lời giải hiện tại hay là nên tạo một lời giải mới. (Việc chọn dựa vào cách mà lời giải đó thỏa yêu cầu từng bài toán, ví dụ như TSP thì khoảng cách sẽ giảm dần, trong khi với bài toán).

## Kết luận chương 2

Như chúng ta đã đề cập phía trên Hill Climbing là một thuật toán metaheuristic, nó không giúp tối ưu hóa bài toán một cách tốt nhất, nhưng nó trả về kết quả làm thỏa mãn yêu cầu bài toán. Có nhiều dạng Hill Climbing nhưng việc áp dụng thuật toán như thế nào cho tốt, chiến lược thế nào cũng một phần góp tích cực cải thiện hiệu suất của bài toán.

Hill Climbing là một trong những thuật toán cùng họ với local search, có rất nhiều thuật toán tương tự nhưng nó là nền tảng bước đệm để chúng ta có thể nghiên cứu sâu hơn nhưng thuật toán khác.

# CHƯƠNG 3: ĐỀ XUẤT THUẬT TOÁN HILL CLIMBING GIẢI BÀI TOÁN NGƯỜI BÁN HÀNG

## Cơ sở lý thuyết

Dựa vào độ phức tạp của bài toán người bán hàng, nếu dữ liệu của thành phố qua nhiều thì bài toán nếu đưa về dạng tìm chính xác kết quả với những thuật toán Vét cạn, đệ quy, quy hoạch động,... Thực sự không khả thi vậy nên cần tìm một lời giải gần đúng ở đây, vậy nên việc áp dụng Hill Climbing vào bài toán là một cách giải quyết vấn đề lớn cho bài toán người bán hàng là hoàn toàn hợp lý. Đáp ứng với yêu cầu và nhu cầu của bài toán.

## Giải bài toán người bán hàng

### Trình bài giải thuật



#### Khởi tạo lời giải ban đầu

Lời giải ban đầu là một yếu tố quan trọng của bài toán, việc tạo ra một lời giải nghĩa là chúng ta đang tạo nên vị trí đứng khác nhau trong bài toán TSP, làm tăng khả năng chúng ta tìm được local maximum hoặc có khi thậm chí là global maximum. Nếu chúng ta chạy liên tục hàm Hill Climbing như vậy thì lời giải cuối cùng của chúng ta nhận được sẽ khác nhau

Chart, diagram

Description automatically generated

Mũi tên đỏ là những vùng chúng ta có thể bước đến ngay ban đầu, từ đó chúng ta dùng giải thuật Hill Climbing để tiếp cận những vùng lân cận và leo đến vị trí cao nhất của vùng đó. Có rất nhiều cách để tạo lời giải ban đầu như là dùng ngẫu nhiên hoặc là dùng heuricstic.

Đối với việc tạo lời giải ban đầu bằng cách ngẫu nhiên, chúng ta đưa bài toán về dạng tạo ngẫu nhiên một mảng n phần tử không trùng nhau, cuối cùng chúng ta thêm thành phố cuối cùng là thành phố ban đầu của chúng ta.

**Sơ đồ khối tạo tuyến đường ngẫu nhiên:**

A picture containing diagram

Description automatically generated

1. Giá trị đầu vào là n độ dài tuyến đường cần tạo ngẫu nhiên.
2. Tạo mảng a có độ dài là n phần tử có giá trị phần tử tăng theo thứ tự. VD: [0, 1 , 2, 3, 4, 5] với n = 6.
3. Khởi tạo mảng kết quả result.
4. Kiểm tra mảng a còn phần tử hay không? Nếu có, chọn ngẫu nhiên một giá trị phần tử trong mảng a ra, đưa giá trị đó vào mảng kết quả result, cuối cùng xóa giá trị đó ra khỏi mảng a. Nếu không, thêm phần tử cuối cùng vào mảng với giá trị là result[0].
5. Kết thúc thuật toán.

**Mô tả bằng code:**



Như vậy chúng ta đã hoàn toàn có thể tạo được hàm khởi tạo lời giải ban đầu cho bài toán.

Chúng ta cần phải có thuật toán tính giá trị của lời giải đó để những bước sau có thể sử dụng. Như chúng ta đã thấy, cứ 2 thành phố sẽ có 1 giá trị trọng số giữa chúng, chúng ta nhóm các cặp thành phố, sau đó cộng tất cả các giá trị lại với nhau.

Ví dụ: chúng ta có tuyến đường [0, 1, 2, 0] thì cách tính thành phố sẽ là:

|  |  |
| --- | --- |
| Thành phố 0 sang thành phố 1 | d1 |
| Thành phố 1 sang thành phố 2 | d2 |
| Thành phố 2 sang thành phố 0 | d3 |

Với d1, d2, d3 là giá trị trọng số từ đỉnh a sang đỉnh b, giá trị đoạn đường . Do đầu vào của bài toán là ma trận tsp (n cạnh x n cạnh) nên việc xác định giá trị d1, d2, d3 hoàn toàn dễ dàng. Nếu từ thành phố 0 sang thành phố 1 thì trong ma trận tsp (n cạnh x n cạnh) giá trị đó sẽ là tsp[0][1], nếu từ thành phố 1 sang thành phố 2 thì giá trị sẽ là tsp[1][2].

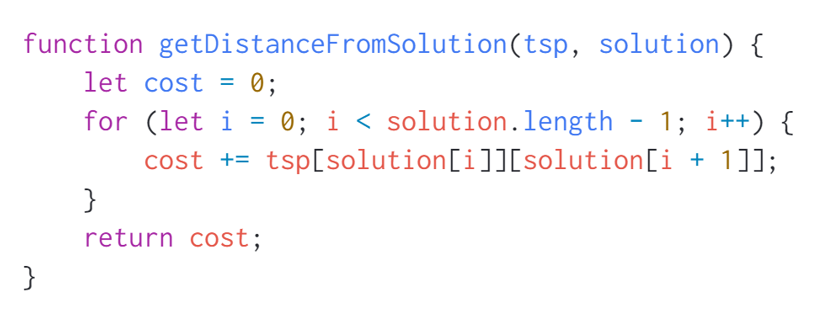
**Sơ đồ khối tính giá trị tuyến đường:**

Diagram

Description automatically generated

1. Nhập vào giá trị tuyến đường đi solution và ma trận tsp.
2. Khởi tạo biến đếm i và giá trị tuyến đường cost là bằng 0.
3. Trong khi i < solution.length – 1, tức là chúng ta chạy đến giá trị kề cuối của solution. Nếu có, cộng vào giá trị đoạn đường vào cost, tăng biến đếm lên 1 . Nếu không kết thúc thuật toán và trả về kết quả.
4. Kết thúc thuật toán.

**Mô tả bằng code:**

****

#### Xác định điều kiện dừng của bài toán

Mấu chốt của bài toán là điều dừng, nếu không xác định điều kiện dừng đúng thì bài toán sẽ lặp vô tận. Ở đây chúng ta muốn tìm những lời giải hàng xóm của lời giải hiện tại, vậy thì bao nhiêu là đủ cho từng bài toán, cho từng trường hợp số thành phố khác nhau. Ý tưởng ở đây là chúng ta sẽ giới hạn mức tìm số lời giải hàng xóm của lời giải hiện tại bằng một con số, nếu như sau số lần tìm mà lời giải không thể cải thiện được hơn nữa thì chúng ta thực hiện dừng bài toán. Vì vậy mà chúng ta sẽ khởi tạo 2 biến (số lần tìm những lời giải hàng xóm xung quanh tại một lời giải hiện tại) và (đếm số lần đã tìm lời giải hàng xóm của lời giải hiện tại). Nếu như sẽ dừng bài toán còn nếu ngược lại thì sẽ thực hiện tìm lời giải hàng xóm và đánh giá kết quả đó là tốt hay là không tốt.

#### Khởi tạo tuyến đường S’ và đánh giá với tuyến đường S

Sau khi xác định điều kiện dừng bài toán, chúng ta bước vào phần chính của giải thuật. Cách tạo S’ lời giải hàng xóm từ tuyến đường S được cho ban đầu. S’ được tạo từ hàm successor function và có rất nhiều cách để tạo:

1. Đảo (chọn 2 điểm, và chuỗi con giữa hai điểm này bị đảo)
2. Chèn (chọn 1 thành phố và chèn nó vào một vị trí ngẫu nhiên)
3. Dời chỗ (chọn một hành trình con rồi chèn nó vào một vị trí ngẫu nhiên)
4. Trao đổi qua lại (hoán vị 2 thành phố)

**Chiến thuật toán đảo:**

|  |  |
| --- | --- |
| **Sơ đồ khối** | **Code** |
| A picture containing text, businesscard  Description automatically generated |  |

**Chiến thuật chèn:**

|  |  |
| --- | --- |
| **Sơ đồ khối** | **Code** |
| **A picture containing text  Description automatically generated** |  |

**Chiến thuật chèn:**

|  |  |
| --- | --- |
| **Sơ đồ khối** | **Code** |
| **A picture containing text  Description automatically generated** |  |

Trên đây là 3 chiến thuật chúng ta áp dụng vào bài toán để sinh ra tuyến đường S’. Sau khi chúng ta xác định được S’ bây giờ sẽ tới bước đánh giá chất lượng lời giải S’.

1. Tạo S’
2. Tính giá trị đường đi S’
3. So sánh giá trị cost của S’ với giá trị cost của S
4. Nếu S’ tốt hơn S thì sẽ thay S thành S’ và sẽ bị gán về thành 0. Vì lúc này bài toán trở thành tìm tuyến đường tốt hơn của tuyến đường S’ vừa mới tìm ra. Ngược lại, không cập nhật tuyến đường S, tăng giá trị lên 1, xác nhận vừa tìm được một lời giải hàng xóm của lời giải S.
5. [Xét lại điều kiện dừng của bài toán](#_Xác_định_điều)

#### Tổng hợp tất cả các bước

Sau khi qua tất cả 3 bước của giải thuật Hill Climbing chúng ta sẽ được lời giải tổng quát như sơ đồ dưới đây.

Diagram

Description automatically generated

### Giải bài toán ví dụ với giải thuật Hill Climbing

Bây giờ chúng ta sẽ trình bài bài toán trên cùng với cách giải thông qua một ví dụ nhỏ với ma trận TSP (5x5). Đây là một trong 20 bộ test được đề cập trong Chương 4.

**Ma trận khoảng cách của bài toán:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **0** | **1** | **2** | **3** | **4** |
| **0** | 0.0 | 3.0 | 4.0 | 2.0 | 7.0 |
| **1** | 3.0 | 0.0 | 4.0 | 6.0 | 3.0 |
| **2** | 4.0 | 4.0 | 0.0 | 5.0 | 8.0 |
| **3** | 2.0 | 6.0 | 5.0 | 0.0 | 6.0 |
| **4** | 7.0 | 3.0 | 8.0 | 6.0 | 0.0 |

**Lời giải:**

|  |  |  |
| --- | --- | --- |
| **Vòng Lặp Thứ** | **Lời Giải** | **Kết Quả Tại Vòng Lặp Thứ I** |
| 1 | **Bước 1:** Chúng ta sẽ khởi tạo giá trị cho bài toán  ,  **Bước 2:** Tạo một lời giải bất kì là:  **Lời giải S**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **0** | **3** | **1** | **2** | **4** |   Chi phí đoạn đường: 27  **Bước 3:** countTotalStepAtCurrent > exceedtime là false  **Bước 4:** Chúng ta sẽ chọn chiến lược đảo, chọn 2 vị trí bất kì của kết quả S đảo chuỗi con giữa 2 đầu mút.  **Lời giải S’**  Với vị trí 1 và 5 ta được:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **0** | **2** | **1** | **3** | **4** |   Chi phí đoạn đường: 27  **Bước 5:** Bây giờ chúng ta sẽ đánh giá S’ tốt hơn S hay không. Do S có khoảng cách hiện tại là 27, sau khi tính S’ chúng ta nhận được khoảng cách vẫn là 27 với kết quả này không làm cải thiện được kết quả. Vậy nên S vẫn giữ nguyên. Tăng countTotalStepAtCurrent lên 1 và quay lại bước 3. | **Đoạn đường**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **0** | **3** | **1** | **2** | **4** |  * Chi phí của tuyến đường này là 27 |
| 2 | **Bước 3:** countTotalStepAtCurrent > exceedtime là false  **Bước 4:**  Với vị trí 0, 4 ta được:  **Đoạn S:**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **0** | **3** | **1** | **2** | **4** |   **Khởi tạo S’:**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **1** | **3** | **0** | **2** | **4** |   Chi phí đoạn đường: 23  **Bước 5:**  Đánh giá S’ có tốt hơn S. Là có 27 > 23. Như vậy thì chúng ta sẽ gán S thành S’, countTotalStepAtCurrent = 0 | **Đoạn đường**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **1** | **3** | **0** | **2** | **4** |  * Chi phí của tuyến đường này là 23 |
| 3 | **Bước 3:** countTotalStepAtCurrent > exceedtime là false  **Bước 4:**  Với vị trí 1, 5 ta được:  **Đoạn S:**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **1** | **3** | **0** | **2** | **4** |   **Khởi tạo S’:**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **1** | **2** | **0** | **3** | **4** |   Chi phí đoạn đường: 19  **Bước 5:**  Đánh giá S’ có tốt hơn S. Là có 23 > 23. Như vậy thì chúng ta sẽ gán S thành S’, countTotalStepAtCurrent = 0 | **Đoạn đường**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **1** | **2** | **0** | **3** | **4** |  * Chi phí của tuyến đường 19 |
| 4 | **Bước 3:** countTotalStepAtCurrent > exceedtime là false  **Bước 4:**  Với vị trí 0, 5 ta được:  **Đoạn S:**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **1** | **2** | **0** | **3** | **4** |   **Khởi tạo S’:**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **3** | **0** | **2** | **1** | **4** |   Chi phí đoạn đường: 19  **Bước 5:**  Đánh giá S’ có tốt hơn S. Là có 19 > 19 là không như vậy vẫn giữ nguyên S, tăng countTotalStepAtCurrent = countTotalStepAtCurrent + 1 | **Đoạn đường**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **1** | **2** | **0** | **3** | **4** |  * Chi phí của tuyến đường 19 |
| 53 | Sau 49 vòng lặp tìm tất cả lời giải xung quanh S.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **1** | **2** | **0** | **3** | **4** |   Chúng ta không tìm thấy được lời giải nào cải thiện được kết quả. Vậy nên, vòng lặp sẽ dừng lại và kết quả cuối cùng chính là S | Vậy kết quả cuối cùng của bài toán chính là:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **4** | **1** | **2** | **0** | **3** | **4** |   Chi phí của đoạn đường là: 19 |

# CHƯƠNG 4 THỰC NGHIỆM ĐÁNH GIÁ

## Môi trường thực nghiệm

### Về máy tính:

|  |  |  |
| --- | --- | --- |
|  | **Bộ phận** | **Mô tả** |
| **1** | Loại máy | HP Pavilion 15 -cs1xxx |
| **2** | CPU | Intel core i7 9nth 8 xung |
| **3** | Hệ điều hành | Windows 10 Pro |
| **4** | Ram | 8GB |
| **5** | Ổ đĩa cứng | 200 SSD + 1TB HDD |

### Về ngôn ngữ:

Ngôn ngữ lập trình Javascipt và NodeJS phiên bản 10x trở lên

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Thư viện** | **Phiên bản** | **Mô tả** |
| **1** | fs | 0.0.1-security | * Đọc, viết, xóa file * Kiểm tra folder, tạo folder |
| **2** | line-by-line | 0.1.6 | * Đọc từng dòng trong file |
| **3** | lodash | 4.17.21 | * Các hàm thư viện toán học và xử lý thông dụng nhanh chóng |

## Hệ thống thực nghiệm

Bộ test được lấy từ trang http://people.sc.fsu.edu và trang http://elib.zib.de. Gồm 20 bộ dữ liệu có kết quả có sẵn và sử dụng thuật toán Hill Climbing để khảo sát 20 bộ.

## Kết quả thực nghiệm

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Stt** | **Tên File** | **Chiều Dài** | **Kết Quả** | **Kết Quả TN** |
| **1** | five\_d | 5 | 19 | 19 |
| **2** | p01\_d | 15 | 291 | 291 |
| **3** | gr17\_d | 17 | 2085 | 2085 |
| **4** | fri26\_d | 26 | 937 | 953 |
| **5** | att48 | 48 | 33522 | 34758 |
| **6** | eil51 | 51 | 426 | 443 |
| **7** | st70 | 70 | 675 | 722 |
| **8** | eil76 | 76 | 538 | 585 |
| **9** | pr76 | 76 | 108159 | 118773 |
| **10** | kroA100 | 100 | 21282 | 22747 |
| **11** | kroC100 | 100 | 20749 | 22587 |
| **12** | kroD100 | 100 | 21294 | 23165 |
| **13** | eil101 | 101 | 629 | 698 |
| **14** | lin105 | 105 | 14379 | 15774 |
| **15** | ch130 | 130 | 6110 | 6860 |
| **16** | ch150 | 150 | 6528 | 7614 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

## Đánh giá chất lượng lời giải và kết quả

## Kết luận và đóng góp.

# TÀI LIỆU THAM KHẢO

1. **Tên tác giả**, tên công trình, nhà xuất bản, năm xuất bản.
2. **Tên tác giả**, tên công trình, nhà xuất bản, năm xuất bản.
3. https://blog.routific.com/travelling-salesman-problem#:~:text=The%20Travelling%20Salesman%20Problem%20(TSP)%20is%20the%20challenge%20of%20finding,computer%20science%20and%20operations%20research.

**https://www.section.io/engineering-education/understanding-hill-climbing-in-ai/**

**Quy định trình bày quyển báo cáo thực tập tốt nghiệp**

Font: Times New Roman

Đóng bìa kiếng, in khổ giấy A4, in 1 mặt

Căn lề:

Lề trái: 3cm

Lề phải:2cm

Lề trên: 2cm

Lề dưới:2cm

Đánh số trang vào cuối giữa trang, trang 1 bắt đầu từ Lời mở đầu. Các trang trước đó đánh i, ii,iii,...trừ trang bìa và trang lót bìa không đánh số trang.

Giãn dòng: Từ 1.3 đến 1.5 lines

Mục lục quyển báo cáo được đánh tự động

Các hình vẽ, bảng biểu được đánh chỉ mục theo mỗi chương và phải có tiêu đề cho các bảng, hình vẽ, các công thức được đánh chỉ số (theo chương) ở bên phải.

Sinh viên chuyển cho 2 cán bộ hướng dẫn file PDF (1 file duy nhất chứa toàn bộ nội dung của quyển báo cáo thực tập tốt nghiệp).

Chương x: size 16, in đậm, chữ hoa

x.1 // size 14, in đậm chữ hoa

x.1.1.// size 13, in đậm

x.1.1.1. // size 13, in nghiêng; đánh chỉ mục tối đa 4 cấp

x.2.

x.2.1.

x.2.1.1.

....