MỤC LỤC

1. Lý do chọn đề tài

Mạng máy tính ngày nay đã trở thành một lĩnh vực phát triển và ứng dụng cốt lõi của công nghệ thông tin. Một trong những tác vụ chủ yếu luôn luôn được thực hiện trên mạng là truyền tải thông tin. Vấn đề làm sao có thể truyền được thông tin đó theo một con đường có chi phí thấp nhất. Và một trong những thuật toán tiêu biểu giải quyết được vấn đề đó là thuật toán Dijkstra.

Thuật toán Dijkstra mang trên nhà toán học người Hà Lan là Edsger Wybe Dijkstra được đề xuất năm 1959. Từ đó trở thành đã có nhiều ứng dụng thực tiễn trong toán học rời rạc và máy tính điện tử. Đồng thời cũng có nhiều phép chứng minh xuất hiện. Tuy nhiên, các chứng minh thường nặng nề về những kiến thức toán học gây khó khăn cho sinh viên trong quá trình tìm hiểu. Vì vậy, cần thiết phải có một phép chứng minh thuật toán này sao cho vừa đảm bảo yêu cầu chặt chẽ về mặt toán học vừa đưa ra cho sinh viên những hiểu biết thật dễ hiểu về mối liên hệ trực tiếp giữa quá trình thiết kế thuật toán với phép chứng minh. Đó cũng là điều mà tác giả của thuật toán Dijkstra đã yêu cầu.

2. Những kiến thức liên quan

2.1. Giới thiệu về lí thuyết đồ thị

2.1.1. Định nghĩa

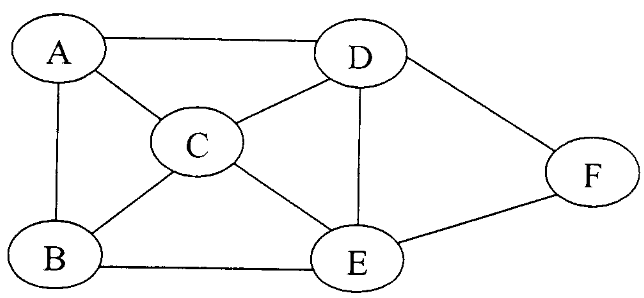
Đồ thị được hiểu là một bộ hai tập hợp hữu hạn: Tập hợp đỉnh và tập hợp các cạnh nối các đỉnh này với nhau.

Ví dụ: G = (V, E)

Trong đó:

V: tập hợp đỉnh

E: tập hợp cạnh (cạnh có hướng hoặc vô hướng nối hai đỉnh (không nhất thiết phải phân biệt))



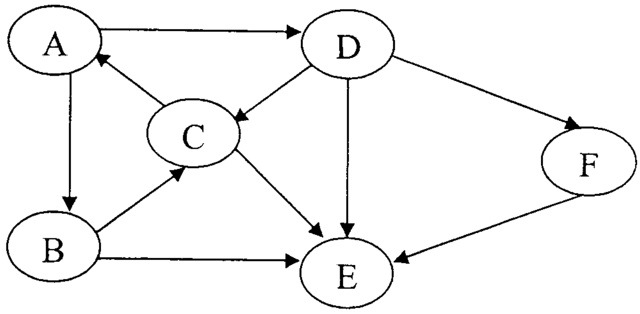
Hình 1. Đồ thị

2.1.2. Phân loại đồ thị

2.1.2.1. Đồ thị vô hướng và đồ thị có hướng

Đồ thị vô hướng G = (V, E): tập cạnh E không phân biệt cạnh nối từ đỉnh X đến đỉnh Y với cạnh nối từ đỉnh Y đến đỉnh X.

Đồ thị có hướng G = (V, E): tập cạnh E phân biệt cạnh nối từ đỉnh X đến đỉnh Y với cạnh nối từ đỉnh Y đến đỉnh X.

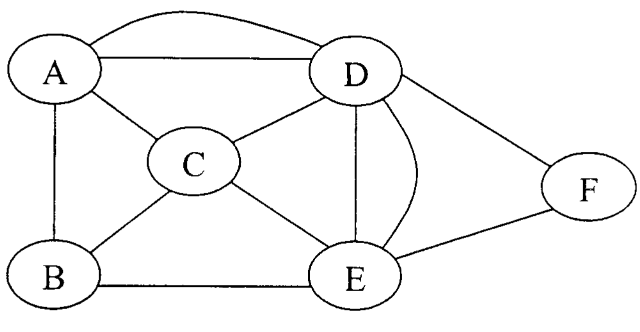


Hình 2. Đồ thị có hướng

2.1.2.2. Đơn đồ thị và đa đồ thị

Đơn đồ thị G = (V, E): tập cạnh E là tập các cặp không có thứ tự gồm hai thành phần tử khác nhau của V gọi là các cạnh.

Đa đồ thị G = (V, E): E là họ các cặp không có thứ tự gồm hai phần tử khác nhau của V gọi là các cạnh. Hai cạnh e_1 và e_2 được gọi là cạnh lặp nếu chúng cùng tương đương với một cặp đỉnh.



Hình 3. Đa đồ thị

2.2. Một số đặc điểm của đồ thị

2.2.1. Đường đi trong đồ thị

Một dãy cạnh $e_1, e_2, \ldots, e_k$ được gọi là dãy cạnh kế tiếp nếu như tồn tại các đỉnh $P_0, P_1, \ldots, P_k$ sao cho $e_i = (P_{i-1}, P_i)$ (cạnh nối P_{i-1} với P_i).

Một dãy cạnh kế tiếp $e_1, e_2,..., e_k$ gọi là đường đi nếu như không có đỉnh nào xuất hiện trên dãy cạnh quá một lần. Nếu $P_0, P_1,...,P_k$ ứng với dãy cạnh $e_1, e_2,..., e_k$ thì $P_i \neq P_j$.

2.2.2. Đồ thị liên thông

Hai đỉnh của đồ thị được gọi là liên thông với nhau nếu chúng có một đường đi từ đỉnh này đến đỉnh kia.

Đồ thị được gọi là đồ thị liên thông nếu hai đỉnh bất kì của nó liên thông với nhau.

2.3. Biểu diễn đồ thị

2.3.1. Biểu diễn đồ thị bằng ma trận kề

Xét đồ thị vô hướng G = (V, E), với tập đỉnh V = {1, 2,..., n}, tập cạnh E = $\{e_1, e_2,..., e_m\}$. Ta gọi ma trận kề của đồ thị G là ma trận A thoả mãn:
$A = \{a_{ij} : i, j = 1, 2,..., n\}$
với các phần tử của A được xác định theo qui tắc sau đây:
$a_{ij} = 0$ nếu $(i, j) \notin E$ và $a_{ij} = 1$ nếu $(i, j) \in E$, với i, j = 1, 2,..., n

2.3.2. Biểu diễn đồ thị bằng ma trận trọng số

Biểu diễn đồ thị bằng ma trận trọng số tương tự như biểu diễn đồ thị bằng ma trận kề, chỉ khác ma trận A được xác định như sau:

Gán a_{ij} giá trị trọng số cạnh nối i với j. Trong trường hợp không có cạnh nối i với j ta gán a_{ij} giá trị $+\infty$.

3. Nội dung chính

3.1. Bài toán tìm đường đi ngắn nhất và thuật toán Dijkstra

Bài toán tìm đường đi ngắn nhất được phát biểu như sau:

Cho đồ thị đơn liên thông G = (V, E) có trọng số không âm. Tìm đường đi với độ dài nhỏ nhất từ một đỉnh xuất phát a tới tất cả các đỉnh còn lại của đồ thị.

Với bài toán trên, ta có thuật toán Dijkstra được viết trên ngôn ngữ tựa Pascal như sau:

Procedure Dijkstra (G: đồ thị đơn liên thông có trọng số không âm);
{ Đánh số các đỉnh của G: $a = v_0, v_1,..., v_n = b$, nối và gán thêm trọng số các cạnh không có trong đồ thị bởi $+\infty$ }
Begin
1. **For** i:= 1 **to** n **do** $L(v_i) := +\infty$;
L(a):= 0;

S:= ∅;
{ Gán nhãn các đỉnh, bắt đầu với nhãn của a là L(a) = 0, còn các đỉnh khác có nhãn bằng +∞. Tập S là tất cả các đỉnh trong quá trình thực hiện được gán nhãn cố định (không cần sửa nhãn nữa đến khi kết thúc thuật toán), các đỉnh còn lại gọi là các đỉnh có nhãn tạm thời. Ban đầu tập S là tập rỗng }

2. **While** <vẫn còn đỉnh có nhãn tạm thời> **do**

Begin

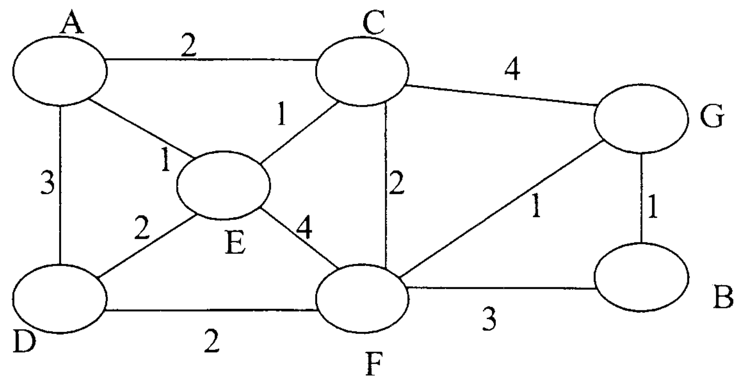
u:= <đỉnh có nhãn tạm thời nhỏ nhất>;
S:= S ∪ {u};
For <tất cả các đỉnh v có nhãn tạm thời> **do**
if L(u) + G(u, v) < L(v) **then** L(v):= L(u) + G(u, v);
{ Thêm vào S đỉnh có nhãn tạm thời nhỏ nhất và sửa đổi nhãn của các đỉnh có nhãn tạm thời còn lại }

End;

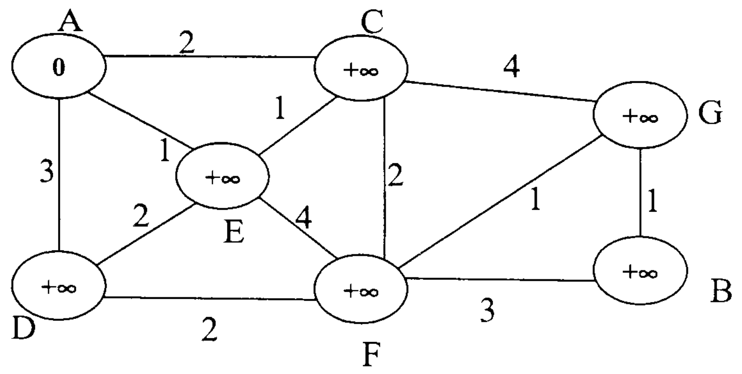
End;

3.2. Ví dụ minh hoạ

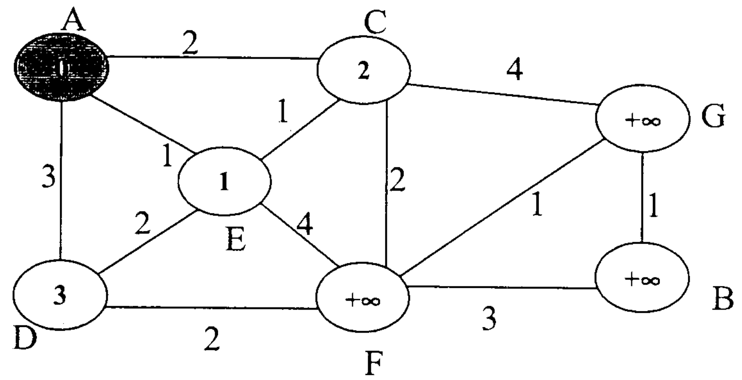
Cho đồ thị sau đây, tìm đường đi ngắn nhất từ a tới b.



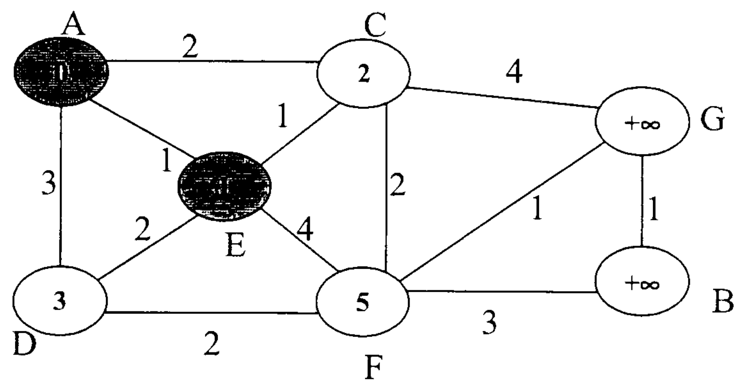
Ta minh hoạ các bước của thực hiện thuật toán bằng đồ thị:



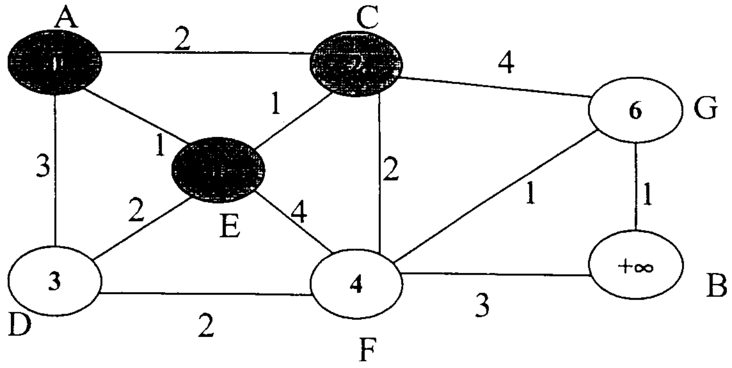
Gán cho A nhãn 0, các nhãn khác có nhãn là +∞



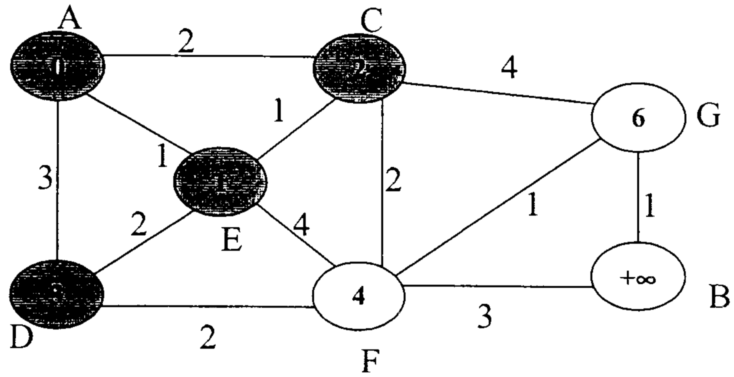
Cố định nhãn đỉnh A, sửa lại nhãn các đỉnh C, D, E



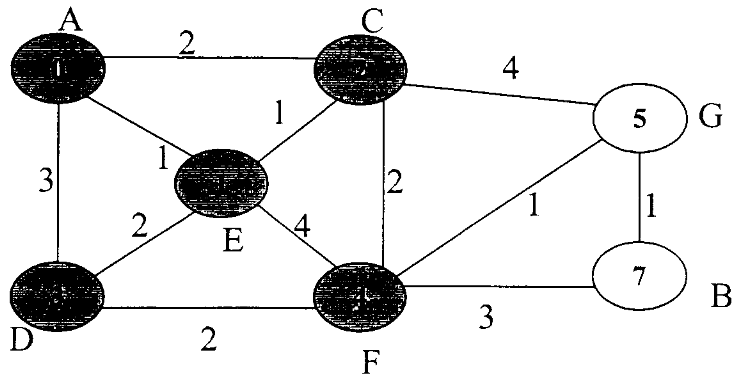
Cố định nhãn đỉnh E, sửa lại nhãn đỉnh F



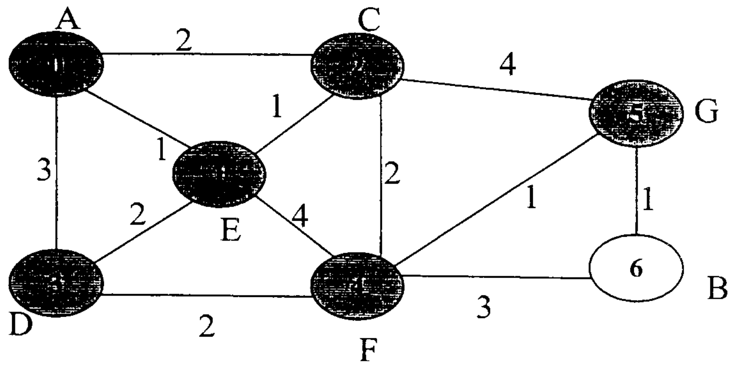
Cố định nhãn đỉnh C, sửa lại nhãn các đỉnh F, G



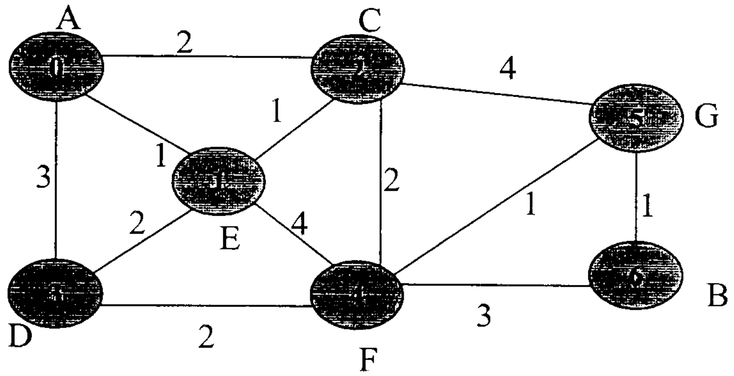
Cố định nhãn đỉnh D, không sửa lại nhãn cho đỉnh nào



Cố định nhãn đỉnh F, sửa nhãn các đỉnh G, B



Cố định nhãn đỉnh G, sửa lại nhãn đỉnh B



Cố định nhãn đỉnh B, kết thúc thuật toán.

Vậy đường đi ngắn nhất từ A tới B có độ dài là 6.

3.3. Phân tích thuật toán

3.3.1. Phép chứng minh trực tiếp

Hiển nhiên thuật toán luôn dừng.

Ta tiến hành phân tích thuật toán thành những bước sau:

Bước 1:

- Gán nhãn cho đỉnh a nhãn $L(a) = 0$, các đỉnh còn lại có nhãn là $+\infty$.
- Tập S các đỉnh có nhãn cố định ban đầu được gán bằng $\varnothing$.

Bước 2:

- Kiểm tra xem còn đỉnh nào mang nhãn tạm thời không. Nếu không thì kết thúc thuật toán.

- Tìm trong số những đỉnh có nhãn tạm thời, đỉnh có nhãn tạm thời nhỏ nhất, giả sử đó là đỉnh x. Cố định nhãn cho đỉnh x (Kết nạp x vào tập S).
- Trong các đỉnh y kề với x và đang có nhãn tạm thời ta thực hiện:

  Nếu $L(y) > L(x) + G(x, y)$ thì gán lại $L(y) = L(x) + G(x, y)$

  Tại các thời điểm mà L(y) được gán nhãn nhãn lại thì giá trị của nó chỉ có giảm đi mà thôi. Vậy khi kết thúc thuật toán, và với tất cả các nhãn x được cố định trước, không tồn tại đỉnh y kề với x, nhưng cố định sau mà $L(y) > L(x) + G(x, y)$. Điều này đồng nghĩa, khi kết thúc thuật toán thì $L(y) \le L(x) + G(x, y)$. Từ đây suy ra khi kết thúc thuật toán thì $G(x, y) \ge L(y) - L(x)$.

Bây giờ ta chỉ xác định đường đi ngắn nhất từ a tới b do việc xác định các đường đi ngắn nhất xuất phát từ a tới các đỉnh khác hoàn toàn tương tự.

Khi thuật toán kết thúc, ta xác định được đỉnh u_1 kề b có tính chất $L(b) = L(u_1) + G(u_1, b)$. Đỉnh u_1 như thế luôn tồn tại vì L(b) giảm dần trong quá trình tính toán còn u_1 chính là đỉnh cuối cùng để giảm L(b). Từ đây suy ra: $G(u_1, b) = L(b) - L(u_1)$.

Tương tự ta cùng tìm được đỉnh u_2 kề u_1 sao cho: $L(u_1) = L(u_2) + G(u_1, u_2)$ ...

Dãy $L(b), L(u_1), L(u_2), \ldots$ giảm dần thực sự, vì vậy, đến một lúc nào đó sẽ có $L(u_{k+1}) = 0$, tức là $u_{k+1} = a$.

Vậy ta có:

$$G(u_1, b) = L(b) - L(u_1)$$
$$G(u_2, u_1) = L(u_1) - L(u_2)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$G(a, u_k) = L(u_k) - L(a)$$

Gọi α là con đường $au_ku_{k-1}\ldots u_2u_1b$.

Gọi $D(\alpha)$ là trọng số đường α. Khi đó:

$D(\alpha) = G(a, u_k) + G(u_k, u_{k-1}) + \ldots. + G(u_2, u_1) + G(u_1, b)$

$= [L(u_k) - L(a)] + [L(u_{k-1}) - L(u_k)] + \ldots + [L(u_1) - L(u_2)] + [L(b) - L(u_1)]$

$= L(b) - L(a) = L(b)$ (Do $L(a) = 0$).

Như vậy trọng số con đường α có giá trị là $D(\alpha) = L(b)$. Ta sẽ chứng minh α là đường đi ngắn nhất nối a với b.

Thật vậy, giả sử β là một con đường bất kì nối a với b. Ta sẽ chứng minh $D(\beta) \ge D(\alpha)$.

Ta đặt $\beta = x_0x_1x_2\ldots x_tx_{t+1}$, các đỉnh x_k và x_{k+1} kề nhau với $0 \le k \le t$, trong đó $x_0 = a$ và $x_{t+1} = b$.

Do khi kết thúc thuật toán một trong hai đỉnh x_k và x_{k+1} phải có đỉnh được cố định trước, một đỉnh được cố định sau. Giả sử đỉnh được cố định trước đó là x, gọi đỉnh còn lại là y.

Vậy:

$L(y) \le L(x) + G(x, y)$

$\Rightarrow G(x, y) \ge L(y) - L(x)$

$\Rightarrow G(x, y) \ge | L(y) - L(x) |$

Từ đây suy ra: $G(x_k, x_{k+1}) = G(x_{k+1}, x_k) \ge | L(x_{k+1}) - L(x_k) |$

Ta có: $D(\beta) = G(a, x_1) + G(x_1, x_2) + \ldots + G(x_t, b)$

$\ge | L(x_1) - L(a) | + | L(x_2) - L(x_1) | + \ldots + | L(b) - L(x_t) |$

$\ge | [L(x_1) - L(a)] + [L(x_2) - L(x_1)] + \ldots + [L(b) - L(x_t)] |$

$= | L(b) - L(a) |$

$= | L(b) |$ (Do $L(a) = 0$)

$= L(b) = D(\alpha)$.

Vậy α chính là đường đi ngắn nhất từ a tới b và theo cách chứng minh L(b) chính là trọng số của đường đi ngắn nhất từ a tới b.

3.3.2. Phép chứng minh qui nạp

Hiển nhiên thuật toán luôn dừng.

Trước tiên, ta chứng minh nhận xét: với mọi $w \in V$ và với mọi thời điểm thực hiện thuật toán ta luôn có: $C(a, w) \le L(w)$, trong đó C(a, w) là độ dài đường đi ngắn nhất từ a tới w.

Thật vậy, ta xét hai trường hợp sau:

- Tại thời điểm đó w chưa được sửa nhãn, vậy thì nhãn của w vẫn là $L(w) = +\infty$. Khi đó hiển nhiên: $C(a, w) \le L(w)$.
- Tại thời điểm đó w đã được sửa nhãn. Theo dõi thuật toán: khi đó L(w) sẽ là độ dài đường đi nào đó nối hai đỉnh a và w. Do C(a, w) là độ dài đường đi ngắn nhất từ a tới w nên $C(a, w) \le L(w)$.

Tóm lại, cả hai trường hợp ta đều thu được: $C(a, w) \le L(w)$.

Từ nhận xét trên ta thu được $C(a, w) \le L^*(w)$, với $L^*(w)$ là giá trị của L(w) khi thuật toán kết thúc.

Giả sử thứ tự các đỉnh đặt vào trong S là: $x_0, x_1, \ldots, x_n$.

Ta sẽ tiếp tục chứng minh rằng: $L^*(x_i) \le L^*(x_{i+1}) \quad \forall i, 0 \le i \le n$.

Khi thực hiện xong vòng lặp thứ i (chọn đỉnh x_i đặt vào S) thì lúc đó $x_{i+1} \notin S$ nên:

$$L^*(x_i) = L_i(x_i) \le L_i(x_{i+1}) \quad (*)$$

($L_i(x)$ là giá trị của L(x) ở vòng lặp thứ i)

Khi thực hiện xong vòng lặp i + 1, ta có:

$L^*(x_{i+1}) = L_{i+1}(x_{i+1})$

$$= \begin{cases} L_i(x_{i+1}) & \text{nếu } L_i(x_{i+1}) \le L^*(x_i) + G(x_i, x_{i+1}) \quad (1) \\ L^*(x_i) + G(x_i, x_{i+1}) & \text{nếu } L_i(x_{i+1}) > L^*(x_i) + G(x_i, x_{i+1}) \quad (2) \end{cases}$$

Với trường hợp (1), theo (*) ta có:

$$L^*(x_{i+1}) = L_i(x_{i+1}) \ge L^*(x_i).$$

Với trường hợp (2) thì hiển nhiên: $L^*(x_{i+1}) > L^*(x_i)$.

Vậy trong cả hai trường hợp (1) và (2) ta đều có $L^*(x_{i+1}) \ge L^*(x_i)$.

Như thế, nếu x được cố định nhãn trước y thì: $L^*(x) \le L^*(y)$.

Trở về vấn đề chính, ta chứng minh khi đỉnh x được gán nhãn cố định thì độ dài đường đi ngắn nhất từ a tới x là $L^*(x)$ bằng phương pháp qui nạp.

- Cơ sở qui nạp: Rõ ràng khi thực hiện khi đỉnh a được cố định nhãn thì $L^*(a) = 0$, hiển nhiên đây là độ dài đường đi ngắn nhất từ a tới a.
- Phần qui nạp:

  Giả sử $L^*[u]$ là độ dài đường đi ngắn nhất từ a tới u với u là đỉnh đang được gán nhãn cố định. Ta sẽ chứng minh đỉnh được gán nhãn cố định tiếp theo trong quá trình thực hiện thuật toán là t sẽ có độ dài đường đi ngắn nhất từ a tới t là $L^*[t]$.

  Trong phép chứng minh trực tiếp ta đã chứng minh được rằng: khi thực hiện xong việc cố định nhãn cho t thì những đỉnh kề t và được cố định nhãn trước ~~x~~ t là x đều thoả mãn $L^*(t) \le L^*(x) + G(x, t)$ và $L^*(x)$ là độ dài đường đi ngắn nhất từ a tới x. Vì vậy, nếu đường đi ngắn nhất từ a tới t có các đỉnh nằm trọn vẹn trong tập các đỉnh đã được cố định nhãn S thì nó phải là $L^*(t)$ vì $L^*(t)$ cũng chính là độ dài một con đường nào đó nối a với t.

  Việc còn lại, chỉ cần chứng minh $L^*(t)$ cũng là nhỏ hơn các đường đi từ a tới t trong đó có những đỉnh không thuộc S.

  Giả sử đỉnh đầu tiên thuộc S trong đường đi từ a tới t là u, đỉnh kề u tiếp theo trong đường đi là v sẽ là đỉnh đầu tiên chưa thuộc S. Vậy đường đi đó được mô tả: λ = a..uv..t. Gọi D(a, t) là trọng số đường đi của λ.

  Do v chưa được cố định nhãn và việc cố định nhãn của v sẽ được thực ở một bước nào đó sau t nên $L^*(t) \le L^*(v)$.

  Cũng theo phần chứng minh trực tiếp ta có: $L^*(v) \le L^*(u) + G(u, v)$ nên:

  $$\begin{aligned} L^*(t) &\le L^*(v) \\ &\le L^*(u) + G(u, v) \\ &\le D(a, u) + G(u, v) \end{aligned}$$

$< D(a, u) + G(u, v) + D(v, t) = D(a, \overset{t}{\cancel{b}}).$

Như vậy, ta đã chứng minh được L*(t) nhỏ hơn hoặc bằng độ dài D(a, t) của một đường đi bất kì từ a tới t. Vậy L*(t) chính là đường đi ngắn nhất từ a tới t.

4. Đánh giá và hướng phát triển

Trong quá trình chứng minh thuật toán Dijkstra bằng hai phép chứng minh: chứng minh trực tiếp và chứng minh qui nạp ta thu được những kết quả sau:

- Với phép chứng minh qui nạp, ta đã chỉ ra, kết thúc thuật toán, đỉnh nào được cố định nhãn trước thì sẽ có nhãn nhỏ hơn hoặc bằng nhãn của đỉnh được cố định sau đó.
- Kết thúc thuật toán, nhãn của các đỉnh là độ dài đường đi ngắn nhất từ đỉnh nguồn tới đó.
- Đường đi ngắn nhất từ đích tới nguồn là đường đi qua các đỉnh được cố định nhãn trước đỉnh nguồn.
- Với phép chứng minh trực tiếp, trong quá trình chứng minh ta đã chỉ ra được rằng đỉnh được sửa nhãn cuối cùng cho đỉnh đích nằm trên đường đi ngắn nhất từ nguồn tới đích. Như vậy, nếu ta lưu đỉnh sửa nhãn cho đỉnh nguồn thì khi kết thúc thuật toán ta tìm được đỉnh sửa nhãn cuối cùng cho đỉnh đích. Cứ như thế lần về tới nguồn. Việc lưu trữ đỉnh trước đó sẽ tăng hiệu quả trong quá trình tìm đường.

Những kết quả này sẽ là tiêu chuẩn đánh giá được sử dụng hầu hết cho các thuật toán tìm đường đi sau này.