

Bài tập 4: Tìm kiếm trên đồ thị lớn

A. Xét tập dữ liệu sgb-words ở địa chỉ:

<http://www-cs-faculty.stanford.edu/~knuth/sgb-words.txt>

Tập dữ liệu này chứa phần lớn các từ tiếng Anh độ dài 5. Từ dữ liệu này, ta xây dựng đồ thị $G = (V, E)$ với tập đỉnh $V =$ “mọi từ trong sgb-words”, và giữa hai từ u và v trong G có cạnh nối nếu u, v khác nhau ở **đúng một** vị trí.

Dễ thấy, trong đồ thị G , dãy từ

words, wolds, golds, goads, grads, grade, grape, graph

là một đường đi với đỉnh bắt đầu là từ words và đỉnh kết thúc là từ graph.

(a) Hãy viết chương trình đếm số thành phần liên thông của đồ thị G .

(b) Hãy viết chương trình nhập vào từ bắt đầu u và từ kết thúc v ; hiện ra màn hình một đường đi ngắn nhất từ u tới v .

B. Vẫn từ tập dữ liệu sgb-words, ta xây dựng đồ thị **có hướng** D với tập đỉnh là “mọi từ trong sgb-words”, và một từ u có cung nối với một từ v khác nếu bốn chữ cuối của u xuất hiện trong v (tính cả số lần xuất hiện).

Đồ thị có hướng D có 94,084 cung và 5757 đỉnh. Đường đi có hướng ngắn nhất từ words tới graph là

words \rightarrow dross \rightarrow soars \rightarrow orcas \rightarrow chars \rightarrow sharp \rightarrow graph

và ta có thể quay lại từ words trong năm bước,

graph \rightarrow harps \rightarrow prats \rightarrow astro \rightarrow trows \rightarrow words

(a) Hãy viết chương trình đếm số thành phần **liên thông mạnh** của đồ thị D .

(b) Hãy viết chương trình nhập vào một từ u và hiện ra màn hình các từ cùng thành phần liên thông mạnh với từ u .

(c) Hãy viết chương trình nhập vào từ bắt đầu u và từ kết thúc v ; hiện ra màn hình một đường đi ngắn nhất từ u tới v .