

Problem: Mining Social-Network Graphs 1

Xét tập dữ liệu sgb-words ở địa chỉ:

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

Tập dữ liệu này chứa mọi từ tiếng Anh độ dài 5. Từ dữ liệu này, ta xây dựng đồ thị $G_{\text{word}} = (V, E)$ với tập đỉnh

$V = \text{“mọi từ trong sgb-words”}$,

và giữa hai từ u và v trong G_{word} có cạnh nối nếu u, v khác nhau ở **đúng một** vị trí.

Để thấy, trong đồ thị G_{word} , 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.

- (1) Hãy viết chương trình đếm số thành phần liên thông của đồ thị G_{word} .
- (2) Hãy viết chương trình nhập vào từ bắt đầu u và từ kết thúc v ; hiện ra màn hình đường đi ngắn nhất từ u tới v .
- (3) Hãy cài đặt thuật toán Girvan-Newman để tìm các cạnh có số betweenness lớn nhất. Hiện các cạnh này ra màn hình.