

Môn: TIN HỌC

Thời gian: 180 phút (không kể thời gian giao đề)

Ngày thi thứ hai: 14/01/2019

(Đề thi có 05 trang, gồm 03 bài)

## TỔNG QUAN NGÀY THI THỨ HAI

	Tên bài	File chương trình	File dữ liệu vào	File kết quả
Bài 4	Nén xâu	COMSTR.*	COMSTR.INP	COMSTR.OUT
Bài 5	Thử nghiệm robot	ROBOT.*	ROBOT.INP	ROBOT.OUT
Bài 6	Thiết kế đường điện	EFILL.*	EFILL.INP	EFILL.OUT

Dấu \* được thay thế bởi PAS hoặc CPP của ngôn ngữ lập trình được sử dụng tương ứng là Pascal hoặc C++.

Hãy lập trình giải các bài toán sau:

## Bài 4. Nén xâu (7 điểm)

Trong quá trình luyện tập cho cuộc thi học sinh giỏi sắp tới, Hùng được thầy giáo giao cho thử sức giải bài toán nén xâu kí tự (chỉ gồm các kí tự la tinh in hoa) sau đây.

Phép cộng trên hai xâu  $x$  và  $y$ , kí hiệu là  $x + y$ , được hiểu là ghép xâu  $y$  liền sau xâu  $x$ . Xuất phát từ hai xâu  $u, v$  và số nguyên  $k$ , xâu  $F_k$  được tạo theo luật sau (còn được gọi là *luật tựa Fibonacci*):

$$F_1 = u; F_2 = v; F_3 = F_2 + F_1; \dots; F_k = F_{k-1} + F_{k-2}.$$

Ví dụ, với hai xâu  $u = 'AB', v = 'C'$  và  $k = 5$  ta có:

$$F_1 = 'AB', F_2 = 'C', F_3 = 'CAB', F_4 = 'CABC', F_5 = 'CABCCAB'.$$

Giả sử xâu  $T$  độ dài  $n$  là xâu được tạo theo luật trên từ hai xâu xuất phát  $u_T, v_T$  có độ dài tương ứng là  $n_1, n_2$ . Như vậy,  $v_T$  là xâu gồm  $n_2$  kí tự đầu tiên của xâu  $T$  và  $u_T$  là xâu gồm  $n_1$  kí tự tiếp theo của xâu  $T$ . Xâu  $T$  có thể được nén thành bộ  $(u_T, v_T, n)$  vì từ 3 thông tin  $u_T, v_T$  và  $n$  ta có thể khôi phục được xâu  $T$  theo luật trên. Ví dụ, xâu  $T = 'CABCCAB'$  có thể được nén thành bộ  $('AB', 'C', 7)$ .

Một xâu  $S$  bất kì có độ dài  $m$  cũng có thể được nén theo cách như trên. Với hai số nguyên dương  $m_1, m_2$  ( $m_1 + m_2 \leq m$ ), gọi  $v_S$  là xâu gồm  $m_2$  kí tự đầu tiên của xâu  $S$  và  $u_S$  là xâu gồm  $m_1$  kí tự tiếp theo trên xâu  $S$ . Khi đó, xâu  $S$  có thể được nén thành bộ  $(u_S, v_S, m)$ . Tuy nhiên, từ 3 thông tin  $u_S, v_S$  và  $m$  ta có thể không khôi phục được chính xác xâu  $S$ . Do đó, người ta đánh giá độ lỗi của phương pháp nén xâu này như sau. Với bộ  $(u_S, v_S, m)$ , tạo xâu  $F_k$  với  $k$  nhỏ nhất mà độ dài  $F_k$  lớn hơn hoặc bằng  $m$  theo luật tựa Fibonacci từ hai xâu xuất phát  $F_1 = u_S, F_2 = v_S$ . Độ lỗi của việc nén xâu  $S$  được tính bằng số lượng vị trí  $i$  mà  $S[i]$  khác với  $F_k[i]$ , trong đó  $S[i]$  và  $F_k[i]$  tương ứng là kí tự thứ  $i$  của xâu  $S$  và xâu  $F_k$  với  $i \leq m$ .

Ví dụ, với  $m_1 = 2$  và  $m_2 = 1$ , xâu  $S = 'CABACC'$  có độ dài  $m = 6$  được nén thành bộ  $('AB', 'C', 6)$ , sau đó tạo ra xâu  $F_5 = 'CABCCAB'$ . Khi đó, độ lỗi của việc nén xâu  $S$  là 2 do có hai kí tự ở các vị trí thứ 4 và thứ 6 của  $S$  và  $F_5$  là khác nhau.