

## §12. KIỂU XÂU

### A. Mục đích, yêu cầu

- Biết xâu là một dãy kí tự (có thể coi xâu là mảng một chiều).
- Biết cách khai báo xâu, truy cập phần tử của xâu.
- Sử dụng được một số thủ tục, hàm thông dụng về xâu.
- Cài đặt được một số chương trình đơn giản có sử dụng xâu.

*Phần minh hoạ với Pascal, học sinh cần nắm được:*

- Cách khai báo (hay định nghĩa) kiểu xâu. Khi khai báo kiểu xâu cần xác định độ dài tối đa của xâu nếu nó nhỏ hơn độ dài được phép (255).
- Có thể nhập hay xuất giá trị của một biến kiểu xâu như đối với một giá trị của biến có kiểu dữ liệu chuẩn và cũng có thể thực hiện các thao tác nhập, xuất hay xử lí từng kí tự của xâu.
- Các phép so sánh (=, <>, <, <=, >, >=) trên xâu được thực hiện theo thứ tự từ điển.
- Có thể ghép các xâu bằng phép ghép xâu (+).
- Ý nghĩa của các hàm chuẩn

`length(s), copy(s, vt, n), pos(s1, s2), upcase(ch)`

và các thủ tục chuẩn

`delete(st, vt, n), insert(s1, s2, vt).`

Học sinh cần có các kĩ năng:

- Khai báo kiểu xâu;
- So sánh hai xâu;
- Nhận biết và bước đầu sử dụng được các hàm, thủ tục chuẩn nói trên.

## B. Những điểm cần lưu ý và gợi ý tổ chức dạy học

### 1. Khó khăn

Một số hàm và thủ tục chuẩn trên kiểu xâu cần được cung cấp cho học sinh, vì học sinh chưa được học về chương trình con, chưa thực sự phân biệt được hàm và thủ tục trong Pascal.

### 2. Thuận lợi

Học sinh đã có kiến thức về kiểu mảng một chiều nên dễ hình dung được cấu trúc và cách sử dụng kiểu xâu.

### 3. Một số lưu ý và gợi ý tổ chức dạy học

Tuy có thể hình dung xâu như mảng một chiều nhưng kiểu mảng một chiều với phần tử thuộc kiểu *char* khác với kiểu xâu (khai báo bằng từ khoá *string*). Lưu ý, không thể áp dụng các thao tác (phép toán, hàm, thủ tục) của xâu cho mảng.

Không nên trình bày sâu về thủ tục và hàm trong bài này. Học sinh đã từng biết và sử dụng một vài hàm chuẩn số học, nên dựa vào đây để giúp học sinh dễ dàng hiểu hàm chuẩn trên kiểu xâu. Hàm chuẩn trên kiểu xâu cho kết quả là một giá trị, tùy theo hàm cụ thể mà giá trị đó có thể là số, kí tự hay xâu.

Ví dụ:

- *length(s)*, *pos(s1, s2)* cho kết quả là số nguyên không âm.
- *copy(s, vt, n)* cho kết quả là một xâu.
- *upcase(c)* cho kết quả là một kí tự, chú ý *c* phải là một phần tử của xâu hay một kí tự.

Có thể chỉ cần giải thích thủ tục chuẩn *delete(st, vt, n)* làm thay đổi xâu *st* (xâu *st* bị xoá bớt một số kí tự) và *insert(s1, s2, vt)* làm biến đổi xâu *s2* (*s2* được thêm một số kí tự).

Cần lưu ý học sinh:

- Xâu được tạo thành bởi các kí tự, trong đó có thể có dấu cách. Dấu cách thể hiện trong các văn bản là phần trống ngăn cách giữa hai từ viết liên tiếp.
- Trong chương trình, khi viết một xâu kí tự, ta phải viết xâu kí tự đó giữa hai dấu nháy đơn. Nhưng khi nhập từ bàn phím giá trị một xâu, ta chỉ cần gõ các kí tự thuộc xâu đó (rồi nhấn phím **Enter**).

- Xâu chỉ gồm một dấu cách được viết là ' '. Để viết xâu rỗng ta viết hai dấu nháy đơn liền nhau.
- Khi so sánh hai xâu, xâu có độ dài nhỏ hơn có thể là xâu lớn hơn, ví dụ:

'Anh' < 'Ba'.

- Khi sử dụng lệnh gán, ta có thể gán trị là một kí tự cho một biến xâu kí tự nhưng việc gán trị là một xâu kí tự cho một biến kiểu kí tự là không hợp lệ dù xâu đó có độ dài bằng 1.
- Tham số của các hàm và thủ tục chuẩn phải hợp lí, chẳng hạn không thể dùng  $insert(s1, s2, 10)$  khi  $length(s2) < 10$ .

Khi giới thiệu hàm và thủ tục thao tác trên xâu, các ví dụ có thể được chuẩn bị trên giấy khổ lớn, treo trước lớp để giáo viên giải thích, tránh mất thời gian viết bảng.

Nên tạo điều kiện cho học sinh vận dụng ngay những hiểu biết mới về các thao tác trên xâu trả lời một số câu hỏi, điều này giúp củng cố và chính xác hoá nhận thức của học sinh. Câu hỏi có thể được đưa ra sau mỗi lần giáo viên giới thiệu một thao tác trên xâu hay sau khi giáo viên giới thiệu hết các thao tác mà bài học đề cập đến. Những câu hỏi như vậy chỉ cần đơn giản kiểu như:

- Giá trị của  $X$  là 'Mon tin hoc', sau thao tác  $delete(X, 4, 4)$  giá trị của  $X$  là gì?
- Giá trị của  $X$  là 'Em yeu truong em', hãy viết thao tác để thêm tên trường (em đang học) vào sau chữ *truong* trong xâu  $X$ .
- Cho biết giá trị của  $pos('xinh', 'Cai xac xinh xinh')$ .

Ở mục 3, qua ví dụ 1 học sinh thấy được cách khai báo và nhập giá trị cho một biến xâu như một biến kiểu đơn giản và cách sử dụng hàm  $length(x)$ . Trong ví dụ này cần lưu ý học sinh phân biệt độ dài tối đa được phép của xâu là 255 và độ dài thực sự của xâu (tùy theo xâu nhập vào cho biến  $a, b$ ).

Ví dụ 2 cho thấy có thể tham chiếu đến từng kí tự của một xâu. Giống như mảng một chiều được đánh chỉ số từ 1, kí tự thứ  $i$  trong xâu được xác định bởi tên biến xâu tiếp đến là  $i$  viết trong hai dấu ngoặc vuông [ và ]. Kí tự cuối trong xâu  $s$  sẽ có vị trí chính bằng giá trị của hàm  $length(s)$ .

Ví dụ 3 củng cố lại việc tham chiếu mỗi kí tự trong một xâu thông qua vị trí của nó.

Ví dụ 4 cung cấp một trường hợp xử lí văn bản đơn giản dùng phép ghép xâu và hàm  $length(s)$  tính độ dài thực sự của một xâu. Trong ví dụ này, một xâu mới

được tạo từ xâu rỗng, qua mỗi bước có thể được ghép thêm một kí tự khác dấu cách.

Giống như ví dụ 4, chương trình ở ví dụ 5 cũng tạo một xâu mới bằng cách ghép dần một số kí tự lấy từ xâu ban đầu, nhưng khác ở chỗ lần này chỉ những chữ số ở xâu ban đầu sẽ được "nhặt ra" để ghép vào xâu mới.

Chú ý rằng các thuật toán ở các ví dụ 3, 4, 5 đều phải thực hiện việc duyệt lần lượt tất cả các kí tự của một xâu nên đều dùng vòng lặp *for-do* (như khi duyệt các phần tử của mảng một chiều).

Tùy theo trình độ học sinh, nên đưa ra các câu hỏi phù hợp để học sinh có thể tham gia xây dựng chương trình ở các ví dụ.

### C. Kiến thức bổ sung

Mỗi biến kiểu xâu với độ dài cực đại  $L$  (được mô tả bởi *string[L]*) dùng  $L+1$  byte để lưu giá trị của xâu mặc dù giá trị thực sự của xâu tại một thời điểm cụ thể có thể nhỏ hơn  $L$ . Trong  $L+1$  byte đó, byte đầu tiên lưu trữ thông tin về số kí tự của giá trị hiện tại (bằng cách chứa kí tự có mã ASCII là độ dài hiện tại của xâu), các byte còn lại thể hiện chính giá trị của xâu.

Ngoài các hàm và thủ tục chuẩn đã trình bày trong sách giáo khoa, còn một số hàm và thủ tục chuẩn sau đây cũng hay được dùng trên kiểu xâu:

- Hàm *concat(S1,S2,...,Sn)*; trả về xâu thu được từ việc ghép các xâu  $S1, S2, \dots, Sn$ .

Ví dụ

Biểu thức	Kết quả
<code>concat('Toi', ' hoc', ' bai')</code>	'Toi hoc bai'

- Thủ tục *str(N,st)*; chuyển đổi giá trị số  $N$  thành kiểu xâu và đưa kết quả vào biến *st*.

Ví dụ

Giá trị N	Biểu thức	Kết quả
1500	<code>str(N, st)</code>	'1500'

- Thủ tục *val(st, N, Code)* chuyển đổi giá trị *st* thành đại lượng kiểu nguyên hay thực và ghi kết quả vào biến  $N$ . Biến *Code* là biến kiểu integer. Nếu việc chuyển đổi không có lỗi thì *Code* sẽ có giá trị 0, trái lại *Code* chứa số chỉ vị trí kí tự đầu tiên phát sinh lỗi và khi đó giá trị của  $N$  là không xác định.

Ví dụ

Giá trị St	Biểu thức	Kết quả
'1500'	val(St, N, Code)	Code = 0, N = 1500
'14.2A+02'	val(St, N, Code)	Code = 5