

§14. KIỂU DỮ LIỆU TỆP

1. Vai trò của kiểu tệp

Tất cả các dữ liệu thuộc các kiểu dữ liệu đã xét đều được lưu trữ ở bộ nhớ trong (RAM) và do đó dữ liệu sẽ bị mất khi tắt máy. Với một số bài toán có khối lượng dữ liệu lớn, có yêu cầu lưu trữ để xử lý nhiều lần, cần có kiểu dữ liệu tệp (file).

Kiểu dữ liệu tệp có những đặc điểm sau:

- Dữ liệu kiểu tệp được lưu trữ lâu dài ở bộ nhớ ngoài (đĩa từ, CD,...) và không bị mất khi tắt nguồn điện;
- Lượng dữ liệu lưu trữ trên tệp có thể rất lớn và chỉ phụ thuộc vào dung lượng đĩa.

2. Phân loại tệp và thao tác với tệp

Xét theo cách tổ chức dữ liệu, có thể phân tệp thành hai loại:

- *Tệp văn bản* là tệp mà dữ liệu được ghi dưới dạng các kí tự theo mã ASCII. Trong tệp văn bản, dãy kí tự kết thúc bởi kí tự xuống dòng hay kí tự kết thúc tệp tạo thành một dòng.

Các dữ liệu dạng văn bản như sách, tài liệu, bài học, giáo án, các chương trình nguồn viết bằng ngôn ngữ bậc cao,... thường được lưu trữ dưới dạng tệp văn bản.

- *Tệp có cấu trúc* là tệp mà các thành phần của nó được tổ chức theo một cấu trúc nhất định. Tệp nhị phân là một trường hợp riêng của tệp có cấu trúc. Dữ liệu ảnh, âm thanh,... thường được lưu trữ dưới dạng tệp có cấu trúc.

Xét theo cách thức truy cập, có thể phân tệp thành hai loại:

- *Tệp truy cập tuần tự* cho phép truy cập đến một dữ liệu nào đó trong tệp chỉ bằng cách bắt đầu từ đầu tệp và đi qua lần lượt tất cả các dữ liệu trước nó.
- *Tệp truy cập trực tiếp* cho phép tham chiếu đến dữ liệu cần truy cập bằng cách xác định *trực tiếp* vị trí (thường là số hiệu) của dữ liệu đó.

Khác với mảng, số lượng phần tử của tệp không cần xác định trước.

Hai thao tác cơ bản đối với tệp là ghi dữ liệu vào tệp và đọc dữ liệu từ tệp.

Thao tác đọc/ghi với tệp được thực hiện với từng phần tử của tệp.

Để có thể thao tác với kiểu tệp, người lập trình cần tìm hiểu cách thức mà ngôn ngữ lập trình cung cấp để:

- Khai báo biến tệp;
- Mở tệp;
- Đọc/ghi dữ liệu;
- Đóng tệp.