



Phân tích thuật toán

Bởi:

Khoa CNTT ĐHSP KT Hưng Yên

Đặt vấn đề

Khi xây dựng được thuật toán để giải bài toán thì có hàng loạt vấn đề được đặt ra để phân tích. Thường là các vấn đề sau:

- Yêu cầu về tính đúng đắn của thuật toán, thuật toán có cho lời giải đúng
- Tính đơn giản của thuật toán. Thường ta mong muốn có được một thuật toán đơn giản, dễ hiểu, dễ lập trình. Đặc biệt là những thuật toán chỉ dùng một vài lần ta cần coi trọng tính chất này, vì công sức và thời gian bỏ ra để xây dựng thuật toán thường lớn hơn rất nhiều so với thời gian thực hiện nó.
- Yêu cầu về không gian : thuật toán được xây dựng có phù hợp với bộ nhớ của máy
- Yêu cầu về thời gian : Thời gian chạy của thuật toán có nhanh không ? Một bài toán thường có nhiều thuật toán để giải, cho nên yêu cầu một thuật toán dẫn nhanh đến kết quả là một đòi hỏi đương nhiên. Trong phần này ta quan tâm chủ yếu đến tốc độ của thuật toán. Ta cũng lưu ý rằng thời gian chạy của thuật toán và dung lượng bộ nhớ nhiều khi không cân đối được để có một giải pháp trọn vẹn. Chẳng hạn, thuật toán sắp xếp trong sẽ có thời gian chạy nhanh hơn vì dữ liệu được lưu trữ trong bộ nhớ trong, và do đó không phù hợp trong trường hợp kích thước dữ liệu lớn. Ngược lại, các thuật toán sắp xếp ngoài phù hợp với kích thước dữ liệu lớn vì dữ liệu được lưu trữ chính ở các thiết bị ngoài, nhưng khi đó tốc độ lại chậm hơn.

Phân tích đánh giá thời gian chạy của thuật toán

- Bước đầu tiên phân tích thời gian chạy của thuật toán là quan tâm đến kích thước dữ liệu như dữ liệu nhập của thuật toán và quyết định phân tích nào là thích hợp. Ta có thể xem thời gian chạy của thuật toán là một hàm theo kích thước của dữ liệu nhập. Nếu gọi n là kích thước của dữ liệu nhập thì thời gian thực hiện T của thuật toán được biểu diễn như một hàm theo n , ký hiệu là: $T(n)$
- Bước thứ hai trong việc phân tích đánh giá thời gian chạy của một thuật toán là nhận ra các thao tác trừu tượng của thuật toán để tách biệt sự phân tích và sự cài đặt. Bởi vì

ta biết rằng tốc độ xử lý của máy tính và các bộ dịch của các ngôn ngữ lập trình cấp cao đều ảnh hưởng đến thời gian chạy của thuật toán, nhưng những yếu tố này ảnh hưởng không đồng đều với các loại máy trên đó cài đặt thuật toán, vì vậy không thể dựa vào chúng để đánh giá thời gian chạy của thuật toán. Ta thấy rằng $T(n)$ không thể được biểu diễn bằng giây, phút... được; Cách tốt hơn là biểu diễn theo số chỉ thị của thuật toán

- Bước thứ ba trong việc phân tích đánh giá thời gian chạy của một thuật toán là phân tích về mặt toán học với mục đích tìm ra các giá trị trung bình và trường hợp xấu nhất cho mỗi đại lượng cơ bản. Chẳng hạn, khi sắp xếp một dãy các phần tử, thời gian chạy của thuật toán hiển nhiên còn phụ thuộc vào tính chất của dữ liệu nhập như:

Dãy có thứ tự đã sắp xếp

Dãy có thứ tự ngược với thứ tự cần sắp xếp

Đã có thứ tự ngẫu nhiên