

II - HƯỚNG DẪN CHI TIẾT

§1. MỘT SỐ KHÁI NIỆM CƠ BẢN

A. Mục đích, yêu cầu

- Biết các vấn đề thường phải giải quyết trong một bài toán quản lí và sự cần thiết phải có CSDL;
- Biết vai trò của CSDL trong học tập và cuộc sống;
- Biết các mức thể hiện của CSDL;
- Biết các yêu cầu cơ bản đối với hệ CSDL.

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

1. Bài toán quản lí

Hệ CSDL được xây dựng và sử dụng trong mọi lĩnh vực hoạt động của con người. Tuy nhiên, chương trình và SGK đã định hướng về hệ CSDL trong quản lí nên GV cần lưu ý làm rõ cho HS thấy các vấn đề sau:

- Công tác quản lí chiếm phần lớn trong các ứng dụng tin học;
- Không phụ thuộc vào lĩnh vực ứng dụng, việc xử lí thông tin trong các bài toán quản lí có những đặc điểm chung: tạo lập hồ sơ, cập nhật hồ sơ, khai thác hồ sơ (sắp xếp, tìm kiếm, tổng hợp thông tin và lập báo cáo...).

Những điều trên được minh hoạ thông qua một ví dụ cụ thể gắn gũi với HS – *Bài toán quản lí học sinh trong nhà trường*. GV cần lưu ý rằng, dù sao đây cũng chỉ là ví dụ nên đã được đơn giản hoá và có nhiều vấn đề phức tạp được bỏ qua.

Cụ thể trong ví dụ này, SGK không đề cập đến số điểm hàng tháng của từng môn, qua đó các tổ bộ môn có thể biết mỗi HS đã có đủ điểm kiểm tra 15 phút, điểm kiểm tra 1 tiết hay chưa,... mà chỉ xét bảng điểm tổng kết một số môn của HS trong một lớp.

Tuy vậy, chỉ với bảng điểm này GV có thể rút ra được nhiều thông tin, trả lời được một số câu hỏi, truy vấn thông thường.

Trước hết, SGK cho thấy việc lưu trữ dữ liệu giúp *trả lời nhanh chóng* được những câu hỏi mà *có thể phải mất rất nhiều thời gian và công sức* để tìm kiếm thống kê trong sổ sách, hồ sơ.

Để HS thấy được vai trò, tác dụng của hệ thống quản lí, GV nên lưu ý phân tích một số vấn đề sau:

- SGK (§ này) đã nêu một số việc thường gặp ở nhà trường, trong đó có việc lưu các thông tin về HS, thống kê kết quả học tập, lập báo cáo tổng kết...
- Ngoài ra, nếu dữ liệu được lưu trữ một cách ***đầy đủ*** và ***hợp lí*** (sẽ nói kĩ ở phần sau) còn giúp "*nhìn thấy*" những điều không ghi trên sổ sách, chẳng hạn như qua điểm số thấy được một HS yếu đang có tiến bộ trong những tuần cuối và cần động viên, tuyên dương kịp thời; các đoàn viên của một chi đoàn nào đó có biểu hiện tư tưởng "trung bình chủ nghĩa" trong học tập ở một số môn không thi đại học,...

Như vậy, dữ liệu không đơn thuần hỗ trợ thống kê, báo cáo, tổng hợp số liệu mà còn giúp người dùng rút ra được kết luận logic cho các hiện tượng diễn ra sau các số liệu tĩnh khô khan trên hệ thống lưu trữ để có những quyết định hợp lí, chính xác trong từng tình huống cụ thể. Tóm lại, nó còn đóng vai trò *hỗ trợ quyết định*.

2. Các công việc thường gặp khi xử lí thông tin của một tổ chức

Dù thông tin được quản lí thuộc lĩnh vực nào, công việc xử lí vẫn phải bao gồm:

- Tạo lập hồ sơ về các đối tượng cần quản lí;
- Cập nhật hồ sơ (thêm, xoá, sửa hồ sơ);
- Khai thác hồ sơ (tìm kiếm, thống kê, sắp xếp, lập báo cáo,...).

Trong SGK không đưa lập kế hoạch và ra quyết định thành một công việc thứ tư cùng với ba công việc trên khi xử lí thông tin mà tách riêng, xem như là mục tiêu chung của ba công việc nêu trên. Tuy nhiên, khi trình bày GV có thể giới thiệu cho HS quan niệm lập kế hoạch, ra quyết định như là công việc thứ tư cũng được.

Nội dung tiểu mục 1 và 2 là những vấn đề liên quan nhiều đến thực tế, không xa lạ với HS. Các khái niệm được trình bày thông qua ví dụ, không đòi hỏi HS phải có sự hiểu biết cặn kẽ. Các thuật ngữ được hiểu tương tự như trong đời thường hoặc đã biết ở lớp 10 và lớp 11 (ví dụ: sắp xếp, tìm kiếm, báo cáo, lập kế hoạch, ra quyết định,...). Không đòi hỏi GV phải trình bày các khái niệm theo cách định nghĩa hình thức, các khái niệm này chủ yếu mang tính giới thiệu ban đầu, nhằm dẫn dắt để sau này HS tiếp thu các kiến thức về hệ CSDL được dễ dàng hơn. Dựa vào ví dụ, GV hướng dẫn, hình thành dần cho HS có hiểu biết, hệ thống hoá các đặc trưng của bài toán quản lí, tạo cơ sở ban đầu để HS dễ tiếp thu kiến thức hơn trong các chương, § tiếp theo.

Có nhiều tình huống GV nên phát vấn HS để dẫn dắt đặt vấn đề. GV nên chuẩn bị sẵn hình 1 SGK trên giấy khổ lớn để không những sử dụng cho việc trình bày §1 mà còn cho nhiều chương, § khác nữa của SGK.

3. Hệ cơ sở dữ liệu

a) Khái niệm cơ sở dữ liệu và hệ quản trị cơ sở dữ liệu

GV nên đưa ra một số câu hỏi hay gặp trong thực tế, ví dụ, qua thông tin có trong hồ sơ lớp: Tổ trưởng quan tâm thông tin gì? Còn lớp trưởng và bí thư chi đoàn muốn biết điều gì?...

Sau khi nghe HS trả lời, GV cần phân tích cho HS thấy:

- Sự đa dạng của các câu hỏi có thể gặp;
- Nguồn gốc của sự đa dạng đó: Có nhiều người cùng khai thác dữ liệu và mỗi người có yêu cầu, nhiệm vụ riêng.

Từ đó, để đáp ứng các yêu cầu khai thác thông tin, phải tổ chức thông tin thành một hệ thống với sự trợ giúp của máy tính điện tử. Với tốc độ truy xuất và xử lí dữ liệu nhanh, máy tính trợ giúp đắc lực cho con người trong việc lưu trữ và khai thác thông tin một cách nhanh chóng và chính xác.

Hoạt động của con người hết sức đa dạng, bao gồm những vấn đề có thể liên quan tới nhiều lĩnh vực như: Kinh doanh, sản xuất, giao thông vận tải, thời tiết,

thể thao, sức khỏe,... Lượng thông tin về mỗi lĩnh vực là rất lớn, vì vậy khó có thể tổ chức một CSDL vạn năng cho tất cả mọi người và đáp ứng mọi yêu cầu. Từ thực tế đó, người ta phải tổ chức nhiều CSDL, mỗi CSDL chỉ liên quan tới một hoặc một số đối tượng và phục vụ cho một người hoặc nhóm người chung mục đích.

Có nhiều định nghĩa khác nhau về CSDL, nhưng các định nghĩa đều phải chứa ba yếu tố cơ bản:

- Dữ liệu về hoạt động của một tổ chức;
- Được lưu trữ ở bộ nhớ ngoài;
- Nhiều người khai thác.

GV có thể đặt câu hỏi cho HS: CSDL của trường và hồ sơ của trường có khác nhau không? Nếu khác thì khác ở những điểm nào? Thông qua câu trả lời của HS, GV nhấn mạnh yếu tố "được lưu trữ ở bộ nhớ ngoài" là rất quan trọng của khái niệm CSDL. Nói một cách khác, yếu tố máy tính ở đây không thể thiếu. Điều này được hệ thống hoá ở trong tiểu mục 3a và đó là nội dung yêu cầu HS phải ghi nhớ.

Yếu tố nhiều người cùng khai thác CSDL và phục vụ nhiều yêu cầu khai thác khác nhau là nhằm đảm bảo hiệu quả khai thác CSDL, giảm giá thành,...

Đôi khi CSDL cá nhân chỉ được xây dựng định hướng cho một người duy nhất khai thác, ví dụ CSDL phục vụ quản lý thư viện cá nhân. Nhưng ngay cả trong trường hợp đó, nó cũng đáp ứng được các yêu cầu của bạn bè và người thân, tức là đáp ứng tính chất nhiều người khai thác.

Một CSDL luôn luôn gắn liền với phần mềm để xây dựng, cập nhật CSDL và khai thác thông tin trong CSDL. Phần mềm này được gọi là hệ QTCSDL.

Hệ QTCSDL bao gồm các môđun chương trình thực hiện các công việc đã nêu. Ngoài ra, nó còn chứa một thành phần hết sức quan trọng là ngôn ngữ giao tiếp. Ngôn ngữ này cho phép người lập trình kích hoạt hay huỷ bỏ hoạt động các môđun trong hệ QTCSDL.

Hệ cơ sở dữ liệu gồm hai thành phần: CSDL và hệ QTCSDL. Tuy vậy, do người dùng thường chỉ quan tâm đến dữ liệu nên người ta thường đồng nhất khái niệm CSDL với hệ QTCSDL và gọi chung cả hai thành phần này là CSDL. CSDL và hệ QTCSDL phải cùng tồn tại và thống nhất với nhau, khi đó mới có thể khai thác thông tin từ CSDL.

GV nên giải thích thêm cho HS về hình 3 SGK, đơn thuần chỉ để minh họa hệ CSDL bao gồm CSDL và hệ QTCSDL, ngoài ra phải có các chương trình ứng dụng để việc khai thác CSDL thuận lợi hơn. Tránh vì cấu trúc của hình vẽ mà HS có thể hiểu nhầm rằng hệ CSDL gồm ba lớp bao nhau.

Ngoài ra, sau khi tìm kiếm, chọn được các thông tin cần thiết vẫn còn phải xử lý tiếp. Các chương trình xử lý này được xây dựng và tạo thành một hệ thống chương trình ứng dụng hỗ trợ sử dụng hệ QTCSDL và khai thác CSDL hiệu quả hơn, phù hợp hơn với yêu cầu thực tế.

b) Các mức thể hiện của cơ sở dữ liệu

Ba mức thể hiện của CSDL được giới thiệu trong SGK thực chất là ba mức trừu tượng hoá của CSDL. Tuy nhiên, ở đây SGK chỉ làm rõ các đối tượng quan hệ với CSDL khác nhau theo mức đòi hỏi sự hiểu biết chi tiết về CSDL.

GV không nên đi sâu vào giải thích tỉ mỉ, chính xác những kiến thức trừu tượng trong mục này. Chỉ yêu cầu học sinh có được những hiểu biết cơ bản sau:

- Sử dụng máy tính, con người tạo lập CSDL và khai thác thông tin trong CSDL một cách hiệu quả. Do vậy, khi nói đến các hệ CSDL một cách đầy đủ nhất sẽ phải nói đến nhiều yếu tố kỹ thuật phức tạp của máy tính. Tuy nhiên, tùy theo mức chuyên sâu của mỗi người trong lĩnh vực công nghệ thông tin hay người dùng CSDL mà có những yêu cầu hiểu biết về CSDL khác nhau. Ba mức hiểu và làm việc với một hệ CSDL là mức vật lý, mức khái niệm và mức khung nhìn.
- Cũng vẫn hệ CSDL đó, nhưng những người dùng đầu cuối được hưởng những ích lợi lớn nhất theo nghĩa họ được tạo một môi trường thuận lợi, thân thiện phù hợp với hiểu biết vốn có của mình, không yêu cầu phải biết những chi tiết phức tạp của chuyên ngành máy tính. Điều này làm cho việc khai thác các ứng dụng CSDL trở nên ngày một phổ biến rộng rãi trong các hoạt động kinh tế xã hội.
- Việc đảm bảo sự tương ứng đúng đắn giữa các mức mô tả và làm việc với một hệ CSDL do các kỹ sư công nghệ thông tin và các phần mềm chuyên dụng chịu trách nhiệm.

GV cần dựa vào CSDL Lớp để giải thích ba mức thể hiện của CSDL.

Khi trình bày về mức khung nhìn, GV cần lưu ý thêm cho HS khung nhìn thực chất là phần CSDL mà mỗi người dùng cần quan tâm. Từ đó HS có thể nhận thức được vì sao chỉ có một CSDL mức vật lí, một CSDL mức khái niệm nhưng có thể có nhiều khung nhìn.

Hình 6 và 7 (SGK) minh họa cho trường hợp cụ thể, với CSDL Lớp nếu bỏ bớt một hoặc một vài thuộc tính nào đó người dùng có thể nhận được một khung nhìn. Chú ý rằng, ngay cả hình thức trình bày những thông tin cho GV chủ nhiệm và GV bộ môn cũng có thể khác nhau. GV chủ nhiệm xem mỗi hồ sơ học sinh trên một trang màn hình (h. 6, SGK), còn GV bộ môn Tin học nhìn thấy kết quả học môn Tin học của nhiều học sinh trên một trang màn hình (h. 7, SGK).

Mặt khác, khung nhìn cũng có thể hiển thị những thông tin kết xuất được từ một vài giá trị nào đó đang được lưu trữ trong CSDL. GV chủ nhiệm cần quan tâm đến điểm trung bình cộng của các điểm trung bình các môn học (h. 6, SGK). GV dạy môn Tin học lại cần biết điểm trung bình cộng của cả lớp và số học sinh trong lớp đạt điểm tổng kết loại giỏi trong môn học này (h. 7, SGK). Những thông tin kết xuất được như vậy chính là kết quả của các hàm tính toán trên các giá trị được lưu trữ ở mức vật lí.

Các mũi tên trên hình 8 SGK mô tả giữa các mức thể hiện của CSDL có một sự tương ứng chặt chẽ, đúng đắn. HS chỉ cần hiểu, giữa các mức mô tả CSDL có tồn tại sự tương ứng nào đó mà không cần phải biết cụ thể đó là các tương ứng nào.

GV cũng nên lưu ý HS, tham số m biểu đạt là không phải chỉ có một mà có thể có nhiều khung nhìn khác nhau, còn tham số n biểu đạt CSDL mức khái niệm không phải chỉ có một mà có thể có nhiều bảng dữ liệu khác nhau (tránh để HS hiểu nhầm là tương ứng với bảng 1 có khung nhìn 1, với bảng 2 có khung nhìn 2,...).

c) Các yêu cầu cơ bản của hệ cơ sở dữ liệu

Ở trên đã giả thiết dữ liệu được lưu trữ **đầy đủ** và **hợp lí**. Vậy thế nào là đầy đủ và hợp lí? Làm thế nào để đảm bảo được điều đó? Để giải quyết những vấn đề này dữ liệu cần phải được thu thập, lưu trữ, truy cập theo những quy tắc nhất định. Ngành chuyên môn hẹp nghiên cứu về cách phân tích, thiết kế, thu thập, lưu trữ, truy cập thông tin thành một hệ thống và xử lí chúng là lí thuyết CSDL.

GV cần lưu ý, SGK chỉ đề cập đến việc tổ chức và khai thác dữ liệu mà không nhằm mục tiêu dạy cho HS về thiết kế và xây dựng công cụ khai thác CSDL. Đó là những vấn đề quan trọng và phức tạp. Ở đây các tác giả chỉ tập trung vào các kết quả đã đạt được *sau khi phân tích thiết kế hệ thống*, từ đó tổ chức lưu trữ dữ liệu thành một hệ thống và tổ chức khai thác chúng.

Về các yêu cầu cơ bản của hệ CSDL, GV cần lưu ý giải thích cho HS những điểm sau:

- *Tính cấu trúc*: Trong các hệ CSDL phổ biến hiện nay, dữ liệu cần được tổ chức dưới dạng các bản ghi (như bản ghi trong PASCAL hoặc C/C++ đã biết). Trong hệ QTCSDL cần phải có các công cụ khai báo cấu trúc, xem, cập nhật, thay đổi cấu trúc. Tuy nhiên, với một CSDL được thiết kế tốt thì việc thay đổi cấu trúc rất ít khi phải thực hiện.
- *Tính toàn vẹn*: Dữ liệu phải thoả mãn một số tính chất nhất định theo *yêu cầu thực tế*. Tuy nhiên, khi đưa dữ liệu vào lưu trữ có thể có sai sót trong khâu ghi chép, thu thập hoặc cập nhật. Hệ thống phải phát hiện và thông báo điều này đồng thời phải có cách *hỗ trợ người dùng để đảm bảo các quy định* trong thực tế. Các công cụ phục vụ việc cập nhật dữ liệu phải kiểm tra dữ liệu để các ràng buộc này luôn được thoả mãn. Điều này giúp cho những người chưa hiểu sâu về chuyên môn nghiệp vụ vẫn có thể làm tốt công việc của mình khi sử dụng hệ CSDL. Ví dụ, một bạn đọc không được mượn quá năm cuốn sách một lần. Nếu bạn đọc nào đó mượn quá năm cuốn sách một lần, thì hệ CSDL sẽ đưa ra thông báo từ chối cho mượn sách.
- *Tính nhất quán*: Tính chất này liên quan tới việc xử lý sai sót khách quan trong quá trình khai thác CSDL. Việc xử lý sai sót rất phức tạp và không phải hệ CSDL nào cũng làm tốt. Giáo viên cần lưu ý HS rằng các hệ CSDL ít hay nhiều phải chú ý đến tính chất này.
- *Tính an toàn và bảo mật thông tin*: GV chỉ cần nhấn mạnh cho HS thấy sự cần thiết của tính chất này, còn cơ chế đảm bảo sẽ được xét kĩ hơn trong chương IV. GV cũng cần lưu ý với HS là phải *bảo vệ nội dung thông tin* (để tránh người không có quyền tác động như xoá, sửa,...) và *bảo vệ giá trị thông tin* (không để rò rỉ thông tin) tức là phải đảm bảo để những người không có quyền biết thì không thể biết được.

- *Tính độc lập*: Khi có một số thay đổi hệ CSDL ở những mức làm việc vật lí hoặc khái niệm nhằm tăng hiệu quả sử dụng hoặc đáp ứng một số yêu cầu phát sinh trong thực tế, người dùng luôn mong muốn hệ CSDL cùng các chương trình ứng dụng đã xây dựng thay đổi ít nhất có thể được (tránh xây dựng lại từ đầu). Đây chính là ý nghĩa của tính độc lập dữ liệu. Các ví dụ trong sách giáo khoa về thay đổi mức vật lí hay mức khái niệm nhằm cho HS thấy được một số trường hợp cụ thể phát sinh nhu cầu thay đổi mức vật lí, mức khái niệm.
- *Tính không dư thừa*: Tính chất này cần được đảm bảo ở khâu phân tích và thiết kế hệ thống. Có thể nêu các ví dụ để thấy thiết kế sai sẽ gây ra những rắc rối, lãng phí như thế nào. Chẳng hạn, thay vì chỉ lưu trữ *ngày sinh*, ai đó có "sáng kiến" lưu trữ thêm *tuổi*. Khi đó hàng năm lại phải cập nhật tuổi cho phù hợp – một công việc mất thời gian và vô nghĩa khi trong CSDL đã có ngày sinh.

Phương pháp giới thiệu các kiến thức phần này chủ yếu cũng nên thông qua ví dụ để HS dễ cảm nhận đúng các yêu cầu của CSDL và điều quan trọng là HS biết được tác dụng và sự cần thiết phải có các yêu cầu đó.

d) Một số ứng dụng

GV có thể liên hệ với bài về ứng dụng Tin học trong SGK Tin học 10 để HS nhớ lại sự đa dạng và tác dụng to lớn của công nghệ thông tin nói chung và của CSDL nói riêng. Tuy nhiên cũng cần nói rõ, trong khuôn khổ bài học này SGK quan tâm chủ yếu là các ứng dụng CSDL trong các công việc quản lí. Không đòi hỏi HS phải liệt kê đầy đủ tất cả các ứng dụng đã nêu trong SGK. GV có thể gợi ý và để HS nêu những ví dụ ứng dụng trong các lĩnh vực khác nhau. Ở mỗi ví dụ, GV cần phân tích cho HS thấy:

- Việc ứng dụng hệ CSDL đã mang lại thay đổi gì;
- Trong mọi hoạt động, con người vẫn đóng vai trò quyết định;
- Có nhiều mức ứng dụng hệ CSDL.