**CÁC VẤN ĐỀ CẦN GIẢI QUYẾT**

1. Hai loại kiểu dữ liệu mới được hỗ trợ trong hệ thống cơ sở dữ liệu đối tượng là gì? Hãy cho một ví dụ về mỗi, và thảo luận về những tình huống ví dụ sẽ bị xử lý nếu chỉ là một RBDMS đã có sẵn?
2. Một người sử dụng phải làm gì để xác định một ADT mới?
3. Một người sử dụng để xác định phương pháp phải những gì có thể làm tăng hiệu quả ? Cho một ví dụ.
4. Điều gì là gắn kết sau cùng của phương pháp? Cho một ví dụ về thừa kế để minh họa sự cần thiết để liên kết động.
5. Hệ thống phân cấp tập hợp là gì? Hãy cho một ví dụ để minh họa hệ thống phân cấp tập hợp tạo điều kiện truy vấn.
6. Thảo luận về cách DBMS khai thác tính đóng gói trong việc thực hiện hỗ trợ cho ADTs.
7. Hãy cho một ví dụ minh họa các hoạt động nesting và unnesting.
8. Mô tả hai đối tượng mà sâu tương đương nhưng không cạn bằng, hoặc giải thích lý do tại sao điều này là không thể.

**CƠ SỞ DỮ LIỆU ĐỐI TƯỢNG**

**Nội dung trình bày:**

* Từ phần 25.1 đến 25.5 tạo động cơ và giới thiệu khái niệm về đối tượng. Các khải niệm được thảo luận trong các phần phổ biến cho cả OODBMSs và ORDBMSs, mặc dù cứ pháp tương tự SQL: 1999. Cụ thể:
  + Phần 25.1: Minh họa lý do tại sao lại mở rộng mô hình quan hệ.
  + Phần 25.2: Thảo luận về kiểu dữ liệu trừu tượng, định nghĩa và thao tác.
  + Phần 25.3: Cách thức và cấu trúc của kiểu dữ liệu trừu tượng.
  + Phần 25.4: Xem xét và nhận dạng đối tượng.
  + Phần 25.5: Sự kế thừa và phân cấp các loại.
* Phần 25.6: Xem xét làm thế nào để tận dụng lợi thế của khái niệm hướng đối tượng mới để thiết kế cơ sở dữ liệu ORDBMS.
* Phần 25.7: Thảo luận một số thách thức mới đặt ra bởi hệ thống quan hệ đối tượng.
* Phần 25.8: Thảo luận về ODL và OQL, các tiểu chuẩn cho OODBMSs.
* Phần 25.9: Trình bày so sánh ngắn gọn của ORDBMSs và OODBMs.

1. **DẪN NHẬP:**

* Với mô hình, hệ thống cơ sở dữ liệu quan hệ hỗ trợ rất tốt các kiểu dữ liệu nhỏ, các tập cố định như (integer, dates, strings), Trong đó có cho thấy hỗ trợ đầy đủ cho các lĩnh vực ứng dụng truyền thống như xử lý dữ liệu hành chính. Trong nhiều lĩnh vực ứng dụng, tuy nhiên, các loại dữ liệu phức tạp hơn nhiều cần được xử lý. Thông thường các loại dữ liệu phức tạp được lưu trữ trong một hệ thốn file lưu trữ. Ví dụ về các lĩnh vực với các dữ liệu phức tạp bao gồm hỗ trợ thiết kế và mô hình hóamáy tính (CAD / CAM), kho lưu trữ đa phương tiện, và quản lý tài liệu.
* Với một lượng lớn dữ liệu ngày càng tăng nhanh chóng, một số tính năng từ DBMS. Ví dụ: giảm thời gian phát triển ứng dụng, đồng thời kiểm soát và phục hồi dữ liệu, hỗ trợ lập chỉ mục, và khả năng truy vấn ngày càng trở nên cần thiết và hấp dẫn. Để hỗ trợ các ứng dụng như vậy, một DBMS phải hỗ trợ các kiểu dữ liệu phức tạp. Khái niệm hướng đối tượng có ảnh hướng mạnh mẻ đến những nổ lực tăng cường hỗ trợ cơ sở liệu phức tạp và dẫn đến sự phát triển của hệ thống cơ sở dữ liệu đối tượng.

1. **CÁC LOẠI HỆ THỐNG CƠ SỞ DỮ LIỆU ĐỐI TƯỢNG:**

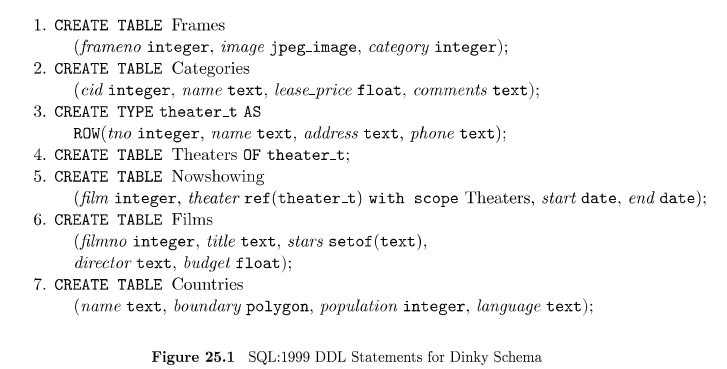
* Hệ thống cơ sở dữ liệu đối tượng có 2 hướng phát triển khác nhau:
  + Hệ thống cơ sở dữ liệu hướng đối đối tượng: hệ thống cơ sở dữ liệu hướng đối tượng là những lĩnh vực nơi các đối tượng phức tạp đóng một vai trò trung tâm. Cách tiếp cận này bị ảnh hưởng nhiều bởi các ngôn ngữ lập trình hướng đối tượng và có thể được hiểu như là một cố gắng để thêm chức năng  DBMS một môi trường ngôn ngữ lập trình.
  + Hệ thống cơ sở dữ liệu quan hệ đối tượng:  Nó có thể được coi là một nỗ lực để mở rộng hệ thống cơ sở dữ liệu quan hệ với các chức năng cần thiết để hỗ trợ một class rộng lớn hơn của các ứng dụng, và bằng nhiều cách, cung cấp một cầu nối giữa các mô hình quan hệ và hướng đối tượng.

🡪Trong seminar này sẽ tập trung vào cơ sở dữ liệu quan hệ đối tượng và nhấn mạnh làm thế nào nó có thể được xem như là một sự phát triển của RDMSs, chứ không phải là một mô hình hoàn toàn khác nhau.

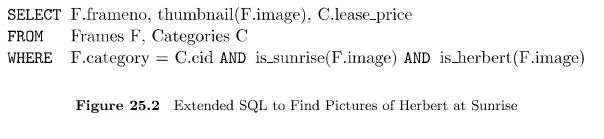
* SQL: tiêu chuẩn năm 1999 là dựa trên mô hình ORDBMS, chứ không phải là mô hình OODBMS. Tiêu chuẩn này bao gồm hỗ trợ cho rất nhiều các tính năng kiểu dữ liệu phức tạp được thảo luận trong chương này. Bài seminar này tập trung vào việc phát triển các khái niệm cơ bản, chứ không phải là trình bày SQL: 1999, một số trong những tính năng mà chúng tôi thảo luận là không bao gồm trong SQL: 1999. Chúng tôi đã cố gắng để phù hợp với SQL: năm 1999 với ký hiệu, mặc dù chúng tôi có đôi khi tách ra một chút cho rõ ràng cho rõ ràng. Điều quan trọng là nhận ra rằng các khái niệm chính được thảo luận phổ biến cho cả hai ORDBMSs và OODBMSs, và chúng tôi thảo luận làm thế nào họ được hỗ trợ trong tiêu chuẩn ODL / OQL được đề xuất cho OODBMSs tại Mục VIII.
* Các nhà cung cấp RDBMS bao gồm cả IBM, Informix, Oracle, có thêm chức năng ORDBMS (mức độ khác nhau) trong các sản phẩm của họ, và điều quan trọng là nhận ra làm thế nào các cơ quan hiện có của kiến thức về thiết kế và thực hiện các cơ sở dữ liệu quan hệ có thể được thừa hưởng để giải quyết các ORDBMS phần mở rộng. Nó cũng rất quan trọng để hiểu những thách thức và cơ hội mà các phần mở rộng trình bày cho người sử dụng cơ sở dữ liệu, thiết kế, và người thực hiện.

1. **Nội dung trình bày:**
2. **Minh họa lý do tại sao lại mở rộng mô hình quan hệ:**

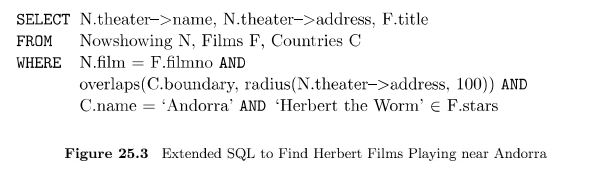
* Như một ví dụ cụ thể về sự cần thiết cho các hệ thống quan hệ đối tượng, chúng tôi tập trung vào một vấn đề kinh doanh xử lý dữ liệu mới khó hơn và (theo quan điểm của tác giả ) giải trí nhiều hơn so với đồng đô la và xu sổ sách kế toán của thập kỷ trước.
* Ngày nay, các công ty trong các ngành công nghiệp như giải trí trong kinh doanh bán dữ liệu số, tài sản cơ bản của công ty của họ không phải là sản phẩm hữu hình, nhưng thay vì các đồ tạo tác phần mềm như video và âm thanh.
* Chúng tôi xem xét Công ty tưởng tượng về Giải trí dinky, một tập đoàn lớn Hollywood có tài sản chính là một bộ sưu tập của các nhân vật hoạt hình, đặc biệt là Herbert âu yếm và quốc tế yêu quý Worm.Dinky có một số bộ phim Herbert Worm, nhiều trong số đó đang được thể hiện trong các rạp chiếu phim trên toàn thế giới tại bất kỳ thời điểm nào. Dinky cũng làm cho một thỏa thuận tốt về tiền bạc cấp giấy phép Herbert của hình ảnh, giọng nói, và đoạn băng video cho các mục đích khác nhau, con số hành động, trò chơi video, xác nhận sản phẩm, và như vậy. Cơ sở dữ liệu của Dinky được sử dụng để quản lý việc bán hàng và hồ sơ cho thuê cho các sản phẩm khác nhau liên quan đến Herbert, cũng như video và âm thanh tạo nên nhiều bộ phim của Herbert.
  1. **Kiểu dữ liệu mới**
* Một vấn đề cơ bản phải đối mặt khi thiết kế cơ sở dữ liệu của Dinky là cần hỗ trợ các loại dữ liệu phong phú hơn đáng kể hơn là có sẳn trong một DBMS:
  + **Người dùng định nghĩa kiểu dữ liệu trừu tượng (ADTs):** Những dữ liệu của Dinky bao gồm hình ảnh, giọng nói, video của Herbert và chúng phải được lưu trữ trong cơ sở dữ liệu. Hơn nữa chúng ta cần các chức năng đặc biệt để thao tác với các đối tượng này. Vi dụ như chúng ta có thể viết các chức năng tạo ra bản nén của một hình ảnh hoặc hình ảnh có độ phân giải thấp hơn.
  + **Kiểu dữ liệu:** Trong các ứng dụng quả thực chúng ta cần xây dựng loại mới dựa vào các hàm dựng như sets, tuples, arrays, sequences.
  + **Tính kế thừa:** Khi số lượng các loại dữ liệu phát triển, điều quan trọng là phải nhận ra sự phổ biến giữa các loại và tận dụng lợi thế của nó. Ví dụ nén hình ảnh và hình ảnh có độ phân giải thấp là 2 loại, nhưng ở mức độ nào đó, chỉ cần 1 loại là đủ. Do đó, mong muốn kế thừa một số tính năng của các đối tượng hình ảnh khi xác định các đối tượng hình ảnh nén và các đối tượng có độ phân giải thấp.
* Làm thế nào chúng ta có thể giải quyết những vấn đề này trong một RDBMS? Chúng ta có thể lưu trữ hình ảnh, video, và như vậy là BLOB trong hệ thống hiện tại quan hệ, một nhị phân lớn đối tượng (BLOB) là chỉ một dòng dài byte, và việc của DBMS hỗ trợ lưu trữ và lấy BLOB trong cách một ví dụ mà người dùng không phải lo lắng về kích thước của BLOB, một BLOB có thể chiều dài một vài trang, không giống như một thuộc tính truyền thống. Tất cả các quá trình gia công, chế biến thêm BLOB đã được thực hiện bằng chương trình ứng dụng của người sử dụng, trong ngôn ngữ máy chủ lưu trữ trong đó các mã SQL được nhúng vào. Giải pháp này không hiệu quả bởi vì chúng ta buộc phải lấy tất cả các BLOB trong một bộ sưu tập ngay cả trong số họ có thể được lọc ra câu trả lời bằng cách áp dụng chức năng người dùng định nghĩa (trong DBMS). Nó là không thỏa đáng từ một quan điểm nhất quán dữ liệu, bởi vì ngữ nghĩa của dữ liệu là phụ thuộc nhiều vào mã ứng dụng ngôn ngữ sở tại và không thể được thi hành bởi DBMS
* Đối với các loại cấu trúc và thừa kế, chỉ đơn giản là không có hỗ trợ trong mô hìnhquan hệ. Chúng ta buộc phải ánh xạ dữ liệu với cấu trúc phức tạp như vậy vào một bộ sưu tập các bảng phẳng. (Chúng ta đã thấy ví dụ của ánh xạ như vậy khi chúng tôi thảo luận về các bản dịch từ sơ đồ ER, thừa kế cho mối quan hệ tại Chương 2.)
* Ứng dụng này rõ ràng đòi hỏi các tính năng không có sẵn trong mô hình quan hệ. Là một minh họa các tính năng, hình 25,1 hiện nay SQL: 1999 DDL báo cáo cho phần của giản đồ ORDBMS Dinky của mà sẽ được sử dụng trong ví dụ tiếp theo Mặc dù DDL là một hệ thống quan hệ truyền thống, nó có một số khác biệt quan trọng mà làm nổi bật việc dữ liệu mới mô hình khả năng của mộtORDBMS.Nhìn lướt qua các câu lệnh DDL là đủ cho bây giờ, chúng tôi sẽnghiên cứu chi tiết trong phần tiếp theo, sau khi trình bày một số các khái niệmcơ bản mà ứng dụng mẫu của chúng tôi cho thấy là cần thiết trong một DBMS thế hệ tiếp theo
  1. **Thao tác với các loại dữ liệu mới**
* Như vậy đến thời điểm này, chúng ta đã mô tả các loại dữ liệu phải được lưu trữ trong cơ sở dữ liệu dinky. Ở đây ta không nói bất cứ điều gì về cách sử dụng các loại mới trong các truy vấn, vì vậy chúng ta hãy nghiên cứu hai truy vấn dữ liệu của dinky cần phải hỗ trợ. Cú pháp của các truy vấn không quan trọng, nó là đủ để hiểu những gì họ thể hiện.Chúng tôi sẽ trở về đặc biệt của cú pháp truy vấn như chúng ta tiến hành



* Thách thức đầu tiên  đến từ công ty Clog cung cấp ngũ cốc ăn sáng. Clog sản xuất ngũ cốc được gọi là Delirios, và muốn thuê một hình ảnh của Herbert Worm ở phía trước của mặt trời mọc, để kết hợp trong thiết kế hộp Delirios. Một truy vấn để trình bày một bộ sưu tập các hình ảnh có thời điểm này có vàgiá cho thuê có thể được thể hiện trong SQL-giống như cú pháp như trong hình25,2. Dinky có một phương pháp số được viết bằng một ngôn ngữ bắt buộc như Java và đăng ký với hệ thống cơ sở dữ liệu. Những phương pháp này có thể được sử dụng trong các truy vấn trong cùng xây dựng trong các phương pháp, chẳng hạn như =, +, -, được sử dụng trong một ngôn ngữ quan hệ như SQL. Các phương pháp hình ảnh thu nhỏ trong mệnh đề Select sản xuất một phiên bản nhỏ của hình ảnh đầu vào đầy đủ kích thước của nó.Các phương pháp is\_sunrise là một chức năng boolean phân tích một hình ảnh và trả về true nếu ảnh có chứa một mặt trời mọc;phương pháp is\_herbert trả về true nếu ảnh có chứa một hình ảnh của Herbert.Truy vấn sản xuất mã số khung, thumhnail hình ảnh, và giá cho tất cả các khung có chứa Herbert và mặt trời mọc.



* Thách thức thứ hai đến từ giám đốc điều hành dinky.Họ biết Delirios là cực kỳphổ biến ở đất nước nhỏ bé của Andorra, vì vậy họ muốn đảm bảo rằng một sốbộ phim Herbert đang chơi tại nhà hát gần Andorra khi cung cấp ngũ cốc việc kệ hàng.Để kiểm tra về hiện trạng của vấn đề, việc giám đốc điều hành muốn tìmtên của tất cả các rạp chiếu phim cho thấy có bộ phim Herbert trong vòng 100km Andorra.Hình 25,3 cho có truy vấn này trong một cú pháp SQL:



* Các thuộc tính rạp chiếu phim của bảng Nowshowing là một tham chiếu đến một đối tượng trong bảng khác, trong đó có các thuộc tính name, address, và location. Tham chiếu đối tượng này cho phép việc ký hiệu N.theater-> name và N.theater-> address, trong đó đề cập các thuộc tính của đối tượng theater\_t tham chiếu có trong hàng Nowshowing N. Thuộc tính có ngôi sao của bảng bộ phim là một tập hợp các tên mỗi ngôi sao của bộ phim. Phương pháp radius trả về một vòng có trung tâm ở số đầu tiên của nó với radius tương đương với số thứ hai của có. Phương pháp overlaps kiểm tra chồng chéo lên nhau không gian. Vì vậy, Nowshowing và Films là sự tham gia của các điều khoản kết bằng, trong khi Nowshowing và các nước tham gia củaviệc khoản chồng chéo không có. Các lựa chọn "Andorra và phim ảnh có chứa ‘Hertbert Worm’ hoàn chỉnh với câu truy vấn.
* Hai câu truy vấn trên tương tự như truy vấn SQL-92 nhưng có một số tính năng không bình thường:
  + Phương thức do người dùng định nghĩa: Người dùng tự định nghĩa các kiểu dữ liệu trừu tượng để thực thi cho các phương thức của họ ví dụ is\_heabert.
  + Các toán tử có cầu trúc: Cùng với các loại cấu trúc có sẵn trong mô hình dữ liệu, ORDBMSs cung cấp việc phương pháp tự nhiên cho các loại dữ liệu đó. Ví dụ, các thiết lập của chuẩn của các loại…
  + Toán tử liên kết: tính tham chiếu trong loại được thế hiện bằng toán tử -> (arrow).

1. **Thảo luận về kiểu dữ liệu trừu tượng, định nghĩa và thao tác**

* Xem xét bảng Frames - 25,1. Nó có một cột hình ảnh của loại jpeg\_image, dùng để lưu trữ một hình ảnh nén đại diện cho một Frame duy nhất của một bộ phim. Các loại jpeg\_image không nằm trong nhân DBMS được xây dựng trong các loại và được xác định bởi một người dùng cho việc ứng dụng dinky để lưu trữ dữ liệu hình ảnhnhất nén bằng cách sử dụng tiêu chuẩn JPEG. Một ví dụ khác, bảng Countries xác định trong dòng 6 của hình 25,1 có một ranh giới cột kiểu của đa giác, trong đó có đại diện của các hình dạng của các Countries vạch ra trên bản đồ thế giới.
* Cho phép người dùng định nghĩa kiểu dữ liệu tùy ý mới là một tính năng quan trọngcủa ORDBMSs. DBMS cho phép người dùng lưu trữ và lấy đối tượng jpeg\_image loại, giống như một đối tượng của bất kỳ loại khác, chẳng hạn như số nguyên. Các kiểu dữ liệu nguyên tử thường cần phải có các loại hoạt động cụ thể quy định bởi người sử dụng tạo ra chúng. Ví dụ, người ta có thể xác định các hoạt động trên một kiểu dữ liệu hình ảnh và các cộng sự của nó phương pháp được gọi là một kiểu dữ liệu trừu tượng, hoặc ADT. Truyền thống SQL đi các với xây dựng trong ADTs, chẳng hạn như số nguyên (với các phương pháp số học có liên quan), hoặc các chuỗi (với sự bình đẳng, so sánh, và phương thức tương tự). Hệ thống quan hệ đối tượng bao gồm các ADTs và cũng cho phép người dùng xác định ADTs của riêng mình.
* Nhãn "abstract" được áp dụng cho các loại dữ liệu vì hệ thống cơ sở dữ liệu không cần phải biết làm thế nào dữ liệu của ADT được lưu trữ cũng như phương thức làm việc của ADT. Nó chỉ đơn thuần là cần phải biết những phương thức có sẵn và các loại đầu vào và ouput cho các phương thức. Cất giấu bên trong ADT được gọi làđóng gói. Lưu ý rằng ngay cả trong một hệ thống quan hệ, các loại nguyên tử chẳng hạn như số nguyên có liên quan đến phương pháp được đóng gói vào ADTs. Trong trường hợp các số nguyên, phương pháp tiêu chuẩn cho ADT là các nhà khai thácsố liệu thông thường và so sánh. Để đánh giá các nhà điều hành bổ liệu vào số nguyên, hệ thống cơ sở dữ liệu không cần phải hiểu những quy định của - luật bổ sung - nó chỉ cần biết làm thế nào để gọi phương thức đoạn mã khai thác và loại dữ liệu để mong đợi trong trả về.
* Trong một hệ thống quan hệ đối tượng, đơn giản hóa do sự đóng gói là rất quan trọng bởi vì nó che giấu bất kỳ sự phân biệt nội dung giữa các loại dữ liệu và cho phép một ORDBMS được thực hiện mà không cần dự đoán các loại và phương thức mà người dùng có thể muốn thêm. Ví dụ, hình ảnh bổ sung thêm các số nguyên và phủ có nhất được xử lý đồng nhất của hệ thống, với sự phân biệt đáng kể duy nhất được rằng mã khác nhau trong gọi cho hai hoạt động và đối tượng khác nhau đã nhập được dự kiến sẽ được trả lại từ mã đó.
  1. **Định nghĩa phướng thức của kiểu dữ liệu trừu tượng**
* Ở mức tối thiểu, đối với từng loại nguyên tử mới, một người sử dụng phải xác định các phương pháp cho phép các DBMS để đọc và để tạo ra các đối tượng của loại hình này và để tính toán lượng lưu trữ cần thiết để giữ đối tượng.Người dùng tạo ra một loại mới nguyên tử phải đăng ký các phương thức sau đây với các DBMS:
  + **Size:** Trả về số byte cần thiết dùng để lưu trữ đối tượng này hoặc một giá trị động đặc biệt, nếu như nó có kích thước động.
  + **Import:** Là phương thức cho phép tạo ra một đối tượng mới dữ liệu đầu vào.
  + **Export:** Ánh xạ các đối tượng để thích hợp cho việc in ấn, hoặc sử dụng trong các chương trình ứng dụng.
* Để đăng ký một phương pháp mới cho một loại nguyên tử, người sử dụng phải viết mã cho phương pháp này và sau đó thông báo cho hệ thống cơ sở dữ liệu về phương pháp.Các mã được viết phụ thuộc vào các ngôn ngữ được hỗ trợ bởi các DBMS, và có thể là hệ điều hành trong câu hỏi.Ví dụ, ORDBMSS có thể xử lý mã Java trong hệ điều hành Linux.Trong trường hợp này, các phương pháp mã phải được viết bằng Java và biên dịch vào một tập tin bytecode Java được lưu trữ trong một hệ thống tập tin Linux.Sau đó, một phong cách phương pháp lệnh SQL đăng ký được trao cho ORDBMS để nó nhận ra phương pháp mới:

1. **Cách thức và cấu trúc của kiểu dữ liệu trừu tượng.**
   1. **Thao tác ở kiểu dữ liệu có cấu trúc**
2. **Xem xét và nhận dạng đối tượng.**
   1. **Khái niệm về Equality**
   2. **Bất tham chiếu đến các loại tham chiếu**
3. **Sự kế thừa và phân cấp các loại.**
   1. **Định nghĩa các loại với sự kế thừa**
   2. **Tính ràng buộc của những phương thức**
   3. **Tập các hệ thống phân cấp, loại mở rộng, và truy vấn**
4. **Xem xét làm thế nào để tận dụng lợi thế của khái niệm hướng đối tượng mới để thiết kế cơ sở dữ liệu ORDBMS**
   1. **Loại có cấu trúc và loại trừu tượng**
   2. **Nhận dạng đối tượng**
   3. **Mở rộng mô hình thực thể kết hợp (ER)**
   4. **Sư dụng tập Nested**
5. **Thảo luận một số thách thức mới đặt ra bởi hệ thống quan hệ đối tượng.**
   1. **Lưu trữ và truy cập phương thức**
   2. **Xử lý truy vấn**
   3. **Tối ưu truy vấn**
6. **Thảo luận về ODL và OQL, các tiểu chuẩn cho OODBMSs.**
   1. **Mô hình dữ liệu ODMG và ODL**
   2. **OQL**
7. **Trình bày so sánh ngắn gọn của ORDBMSs và OODBMs**
   1. **So sánh RDBMS và ORDBMS**
   2. **So sánh OODBMS và ORDBMS: Sự giống nhau**
   3. **So sánh OODBMS và ORDBMS: Sự khác nhau**