

مدل داده

مدل کردن داده روشی برای توصیف داده ها و عملیات روی آنها در سطوح مختلف معماری پایگاه داده است. مدل های سلسله مراتبی و شبکه از قدیمی ترین ساختارها برای مدل کردن داده بوده اند.

مدل کردن داده (data modeling) تحلیل و توصیف داده های محیط عملیاتی و ارتباط بین آنها و شرح معنی و قیدهای داده است.

یک مدل داده قالب قراردادی برای ساخت و کارکردن با داده در اختیار می گذارد.

یک مدل داده پدیده های دنیای واقعی را توصیف می کند و ساختار داده مورد نیاز برای حل مسئله خاص و ارتباط درونی بین ساختارها را تعریف می کند. دیاگرام ER مثالی از این نوع است که بیشتر توسط طراحان و تحلیل گران سیستم به کار می رود.

علاوه بر این مدل داده تعیین می کند یک DBMS چگونه داده را درون خود، به کاربران و برنامه های کاربردی نمایش دهد. مدل رابطه ای مثالی برای این نوع از مدل داده است.

عناصر مدل داده

هر مدل داده باید جنبه های زیر را دارا باشد و نمادهائی برای بیان آنها داشته باشد:

- ساختمان های داده. که راهی برای سازماندهی داده های محیط عملیاتی است و به DBMS می گوید چگونه داده نمایش داده می شود.
- جامعیت. برای تضمین سازگاری و اعتبار داده در پایگاه داده است و به DBMS می گوید چگونه مانع ورود داده نامعتبر به پایگاه داده بشود.
- عملیات. مجموعه روش هایی که امکان کار با داده را می دهند مانند عملیات اضافه، حذف، اصلاح و بازیابی داده.

ساختمان های داده

لازمه ی هر مدل داده ای وجود يك ساختار داده ای است. ساختار های داده نحوه سازماندهی داده های محیط عملیاتی و نمایش داده در DBMS را تعیین می کند.

هر مدل داده باید شامل نمادهائی برای تعریف انواع موجودیت، انواع صفات خاصه و انواع ارتباط بین موجودیت ها باشد.

جامعیت

در هر مدل داده باید امکان تعریف قیدهای جامعیت وجود داشته باشد. قیدها قوانینی را برای تضمین سازگاری و اعتبار داده در پایگاه داده است وضع می کنند و به DBMS می گویند چگونه مانع ورود داده نامعتبر به پایگاه داده بشود.

عملیات

یک مدل داده علاوه بر ساختمان داده، مجموعه عملیاتی را تعریف می کند که می تواند روی داده انجام شود.

عملیات امکان کار با داده را می دهند مانند اضافه، حذف، اصلاح و بازیابی داده.

البته این عملیات ممکن است صریحا در زبان پرس و جو خاصی نباشند. بلکه اساسی را فراهم کنند که بر پایه آن زبان پرس و جو ساخته می شود.

انواع مدل های داده

- مدل موجودیت-ارتباط
- مدل سلسله مراتبی
- مدل شبکه ای
- مدل های داده شیء گرا
- مدل داده نیمه ساختیافته (XML)
- مدل رابطه ای

مدل سلسله مراتبی

قدیمی ترین مدل برای طراحی پایگاه داده مدل سلسله مراتبی (hierarchical model) است، که در اوایل دهه ۶۰ توسط IBM برای سازماندهی دنیای تجارت به شکل سلسله مراتبی پیشنهاد شد.

در مدل سلسله مراتبی داده ها و ارتباط بین آنها به کمک يك درختواره نمایش داده می شوند.

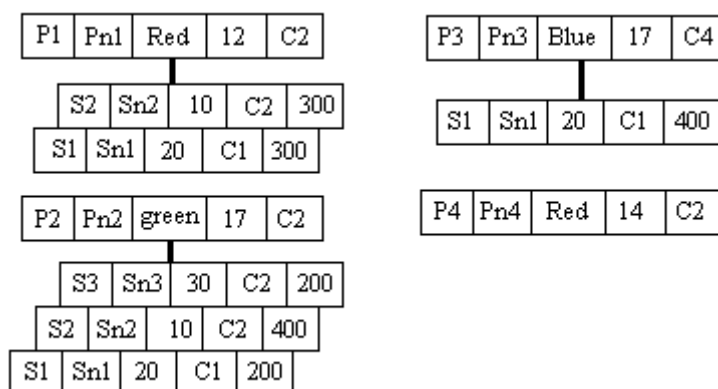
درختواره گرافی است دارای يك ریشه و تعدادی گره که بین هر دو گره پیوندی وجود دارد. پیوند يك مسیر منطقی را از یک گره در سطح بالا به یک گره دیگر در سطح پایین برقرار می کند. مسیری از گره سطح پایین تر به گره سطح بالاتر وجود ندارد.

مدل اجازه تکرار اطلاعات را توسط ارتباطات والد/فرزند می دهد؛ یعنی هر گره در هر سطح می تواند تعدادی گره وابسته یا فرزند داشته باشد که بعنوان والد آنها محسوب می شود. هر گره فرزند تنها دارای یک گره والد است.

مثال. شرکتی چند شعبه دارد. هر شعبه احتیاج به چندین قطعه برای ساخت محصولاتش دارد. هر قطعه از چند تهیه کننده تهیه می شود.

در ساختار سلسله مراتبی دو نوع رکورد قطعه و موجودیت به صورت زیر تعریف می شوند:

Product (P#, Pname, Color, Weight, City)
Supplier (S#, Sname, Status, City, QTY)



خواص مدل

- هر گره درختواره حاوی کلیه صفات خاصه یک نوع رکورد تحت یک نوع موجودیت است.
- درختواره مجموعه ای از رکوردها با انواع مختلف می تواند باشد.
- مجموعه ای از پیوندها را دارد که کلیه انواع رکورد را در ساختار بهم متصل می کند.
- حداکثر یک پیوند بین دو نوع رکورد وجود دارد بنابراین پیوندها نامگذاری نمی شوند.
- هر رکورد تنها والد در سطح بالاتر در درخت وجود دارد.
- اتصال بین رکوردهای هم نوع وجود ندارد. بین رکوردهای هم سطح نمی توان حرکت کرد مگر اینکه والدشان یکی باشد.
- نقطه ورود به ساختار همیشه ریشه است و مسیر منطقی همیشه از بالا به پایین است.
- دو عملگر جداگانه برای یافتن داده ای در ریشه و پرس و جو در فرزندان مورد نیاز است.

مزایا و معایب

- چون داده به صورت یک درختواره سازماندهی می شود برای داده هایی که ماهیت سلسله مراتبی دارند مناسب است.
- ساختار سلسله مراتبی برای مدل کردن ارتباطات یک به چند مناسب است.
- ساختار درختی انعطاف پذیر نبوده و برقراری پیوندهای جدید بین رکوردها بعد از ایجاد درخت امکان ندارد. علاوه بر این پیوندهای عمودی هستند و نمی توان پیوند افقی و مورب در درخت تعریف کرد.
- ارتباط تنها می تواند به صورت "تعلق دارد" یا "شامل می شود" کد شوند.
- در عملیات ذخیره سازی یعنی درج، حذف و بهنگام سازی دارای آنومالی است (درمثال امکان ذخیره تهیه کننده S4 تا زمانیکه ندانیم چه قطعه ای تولید می کند وجود ندارد).
- رویه های پاسخ به پرس و جوهای گزینه متقارن نیستند (برای مثال رویه جستجو برای "شماره تهیه کنندگانی که P2 را تهیه می کنند" متفاوت از رویه جستجوی "شماره قطعاتی که توسط S1 تهیه شده است").
- در نمایش افزونگی داده ناتوان است (در مثال اگر بخواهیم شهر S1 را تغییر دهیم در همه نمونه رکوردهایی که S1 ظاهر شده باید انجام شود غیر اینصورت ناسازگاری رخ می دهد).

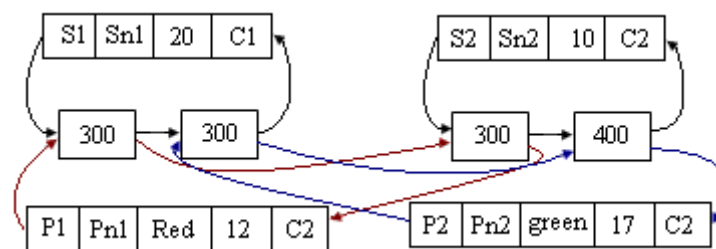
مدل شبکه ای

درمقایسه با مدل سلسله مراتبی که ساختمان های داده ای به صورت درختی از رکوردها سازماندهی می شود و هر رکورد آن یک والد و چند فرزند دارد، مدل شبکه اجازه رکوردهائی با چند والد و چند فرزند را می دهد که در نتیجه یک ساختار مشبک را می سازد. به مدل شبکه ساختار Plex هم گفته می شد.

درمقایسه با مدل سلسله مراتبی که درختی از رکوردها سازماندهی می شود و هر رکورد آن یک والد و چند فرزند دارد، مدل شبکه اجازه رکوردهائی با چند والد و چند فرزند را می دهد که در نتیجه یک ساختار مشبک را می سازد.

عملیات در مدل شبکه به صورت پیمایشی است از یک رکورد به دیگری با دنبال کردن ارتباطاتی که رکورد در آنها سهمیم است دنبال می شود.

مثال. ارتباط دو سویه قطعه و تهیه کننده را در نظر بگیرید. هر قطعه توسط چند تهیه کننده تهیه می شود و هر تهیه کننده چند قطعه را عرضه می کند.



خواص مدل

- در این ساختار موجودیت ها به کمک انواع رکوردها، و ارتباطات به کمک پیوندهای بین رکوردها نمایش داده می شوند.
- هر گره فرزند می تواند بیش از یک گره والد داشته باشد.
- برای نمایش ارتباطات یک به چند دو سویه مناسب است.
- عملیات ذخیره و بازیابی پیچیده تر از مدل سلسله مراتبی است.
- برای پرس و جوهای قرینه رویه پاسخگوئی قرینه دارد ولی پیچیده است.
- متدهائی را برای ساخت و تعریف دوباره پیوندها دارد.
- ساختارهای فیزیکی مناسب برای ارتباطات شبکه ای ساختار چندحلقه است.

مزایا و معایب

- اجازه مدل کردن طبیعی تر ارتباطات مابین موجودیت ها را می دهد.
- مدل شبکه انعطاف پذیری بیشتری نسبت به سلسله مراتبی دارد.
- در عملیات ذخیره سازی انومالی ندارد.

- مدل شبکه می تواند افزونگی داده را بهتر از سلسله مراتبی نشان بدهد.
- از دید کاربر وضوح کامل ندارد
- عملیات پیچیده تری دارد
- پیوند بین رکوردهای یک نوع ممکن نیست.
- اصل وحدت عملگر در یک عمل رعایت نمی شود.

پایگاه داده XML

یک پایگاه داده XML سیستم نرم افزاری است که اجازه می دهد داده در فرمت XML وارد، پردازش و ارسال شود.

دو دسته اصلی پایگاه داده XML وجود دارد:

۱. XML-enabled . پایگاه داده ای که مستند XML را به عنوان ورودی گرفته به یک پایگاه داده دیگر نظیر رابطه ای تبدیل می کند و پس از انجام عملیات آنرا مجدداً به XML بر می گرداند.
۲. Native XML (NXD). مدل داخلی چنین پایگاه داده ای بر پایه XML است و مستندات XML را به عنوان منبع ذخیره سازی مستقیماً استفاده می کند.

دلیل استفاده XML در پایگاه داده شفافیت داده است. داده از پایگاه داده استخراج می شود و در مستندات XML قرار می گیرد و برعکس. به این صورت هزینه ذخیره داده در فرمت XML هم کمتر می شود.

مدل رابطه ای

در این مدل داده ها به شکل جداول سازماندهی می شوند بطوری که ما بین جداول ارتباطات ضمنی نیز وجود دارد، این مدل هیچ یک از مشکلات مدل سلسله مراتبی را ندارد و تنها ایرادی که می توان بر آن گرفت آن است که این مدل فاقد توانایی ذخیره سازی هر گونه داده ای می باشد.

مدل شیء گرا

در این مدل پایگاه داده سعی بر آن است که از مفاهیم شیء گرایی مانند: وراثت، چندریختی و غیره نشود و هنوز یک Data Base کاملاً شیء گرا طراحی نشده است .

تمرین:

۱- کدام گزینه از مزایای یک سیستم پایگاه داده نیست؟

- ☐ ۱. کاهش افزونگی
- ☐ ۲. سازگاری
- ☐ ۳. پیچیدگی
- ☐ ۴. محدودیت های امنیتی

۲- کدامیک از مشکلات زیر ممکن است به دلیل استفاده از قفل توسط تراکنش رخ بدهد؟

- ☐ ۱. رونویسی اطلاعات
- ☐ ۲. از بین رفتن اطلاعات
- ☐ ۳. بن بست
- ☐ ۴. از بین رفتن جامعیت

- ☐ ۱. از دید کاربر ساختار سلسله مراتبی نسبت به ساختار شبکه دارای وضوح بیشتری است.
- ☐ ۲. در مدل شبکه هر فرزند می تواند بیش از یک پدر داشته باشد.
- ☐ ۳. عملیات بازیابی در مدل شبکه از مدل سلسله مراتبی ساده تر است.
- ☐ ۴. برای ارتباطات یک به چند یکسویه بین انواع موجودیت ها مدل سلسله مراتبی مناسب تر است.