



مؤسسه آموزش عالی بهار

غیرانتفاعی - غیردولتی



پایگاه داده ها

تهیه کننده: فرزاد زندی

zandi_farzad@yahoo.com

zandi8farzad@gmail.com

مفاهیم بنیادی پایگاه داده‌ها و عناصر اصلی محیط آن

آنچه در این جلسه می خوانید:

۱- داده

۲- اطلاع

۳- دانش

۴- پایگاه داده‌ها چیست

۵- عناصر محیط پایگاه داده‌ها

۶- انواع سخت‌افزارهای محیط پایگاه داده

۷- انواع نرم‌افزارهای موجود در محیط پایگاه داده‌ها

مقدمه:

اصطلاح پایگاه داده‌ها یکی از رایج‌ترین اصطلاحات در دانش و فن کامپیوتر است

در این درس دانشجویان تنها با بخشی از مفاهیم بنیادی دانش و فن پایگاه داده‌ها آشنا شده، آگاهی پایه‌ای لازم را برای مطالعه بیشتر و یا کار در این زمینه کسب می‌کنند.

مقدمه:

سیستم مدیریت پایگاه داده‌ها یکی از سیستم‌های ذخیره و بازیابی اطلاعات است.

سیستم ذخیره و بازیابی اطلاعات در معنای عام:

هر سیستمی که به کاربر برنامه‌ساز یا نابرنامه‌ساز امکان دهد تا اطلاعات خود را ذخیره، بازیابی و پردازش کند.

داده, اطلاع و دانش:

■ داده:

نمایش ذخیره شده اشیاء فیزیکی، چیزهای مجرد، بوده ها، رویدادها یا چیزهای قابل مشاهده که در تصمیم سازی بکار می آیند.

■ اطلاع:

اطلاع، داده پردازش شده است.

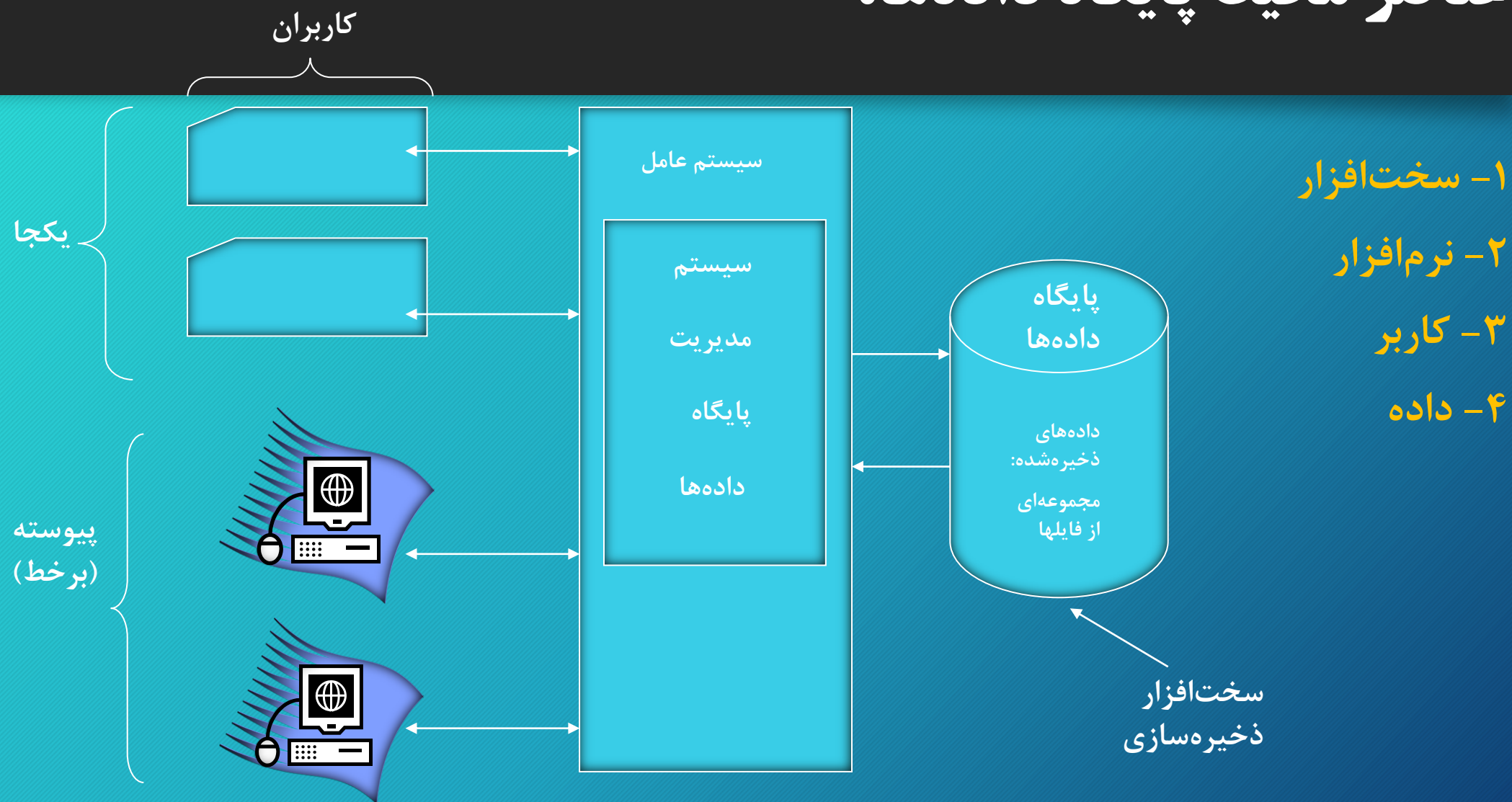
■ دانش:

دانش عبارتست از نمایش نمادین جنبه هایی از بخشی از خرد جهان واقع.

پایگاه داده‌ها چیست؟

پایگاه داده‌ها مجموعه‌ای است از داده‌های ذخیره شده و پایا، به صورت مجتمع (یکپارچه، نه لزوماً فیزیکی، بلکه حداقل به طور منطقی)، بهم مرتبط، با کمترین افزونگی، تحت مدیریت یک سیستم کنترل متمرکز، مورد استفاده یک یا چند کاربر از یک یا بیش از یک "سیستم کاربردی"، به طور همزمان و اشتراکی می‌باشد.

عناصر محیط پایگاه داده‌ها:



انواع سخت افزارهای محیط پایگاه داده:

- سخت افزار ذخیره سازی داده ها

- سخت افزار پردازشگر

- سخت افزار هم‌رسانی (ارتباط)

انواع نرم افزارهای موجود در محیط پایگاه داده‌ها:

- سیستم مدیریت پایگاه داده‌ها (DBMS)

- برنامه‌های کاربردی قابل اجرا در محیط DBMS

- رویه‌های ذخیره‌شده

- نرم افزار شبکه

جلسه دوم

مدلسازی معنایی داده‌ها

آنچه در این جلسه می خوانید:

- ۱- مدلسازی معنایی داده‌ها
- ۲- انواع روشهای مدلسازی معنایی داده‌ها
- ۳- مدلسازی با روش ER
- ۴- تعریف موجودیت
- ۵- قوانین تشخیص موجودیت
- ۶- موجودیت مستقل و موجودیت وابسته
- ۷- تعریف صفت
- ۸- انواع صفت
- ۹- صفت ساده, مرکب, تک مقداری و چند مقداری
- ۱۰- صفت شناسه یا ناشناسه, هیچ مقدارپذیر یا هیچ مقدارناپذیر, ذخیره شده (واقعی) یا مشتق
- ۱۱- تعریف ارتباط

مدلسازی معنایی داده‌ها:

داده‌های ذخیره‌شده در پایگاه داده‌ها ابتدا باید در بالاترین سطح انتزاع مدلسازی معنایی شوند.

مدلسازی معنایی داده‌ها یعنی ارائه مدلی از داده‌های محیط عملیاتی به کمک مجموعه‌ای از مفاهیم انتزاعی و مستقل از جنبه‌های مربوط به نمایش منطقی و فیزیکی داده‌ها، با توجه به معنایی که کاربر برای داده‌ها قائل است.

انواع روشهای مدلسازی معنایی داده‌ها:

- روش موجودیت – ارتباط (ER: Entity Relationship)
- روش زبان عمومی مدلسازی (UML: Unified Modeling Language)
- روش تکنیک مدلسازی شیء (OMT: Object Modeling Technique)

مدلسازی با روش ER:

در روش ER سه مفهوم معنایی وجود دارد و داده‌ها به کمک همین سه مفهوم نمایش داده می‌شوند:

- موجودیت

- صفت (خصیصه)

- ارتباط

تعریف موجودیت:

موجودیت , مفهوم کلی شیء، چیز، پدیده و به طور کلی هر آنچه که می‌خواهیم در موردش اطلاع داشته باشیم و شناخت خود را در موردش افزایش دهیم.

مثال:

محیط عملیاتی: دانشکده

موجودیت: دانشجو , درس , استاد و ...

قوانین تشخیص موجودیت:

- موجودیت معمولاً نمونه‌هایی متمایز از یکدیگر دارند.
- موجودیت معمولاً بیش از یک صفت دارد.
- موجودیت معمولاً یک نام با معنای مشخص دارد.
- موجودیت معمولاً حالت کنشگری (فاعلیت) یا حالت کنشپذیری (مفعولیت) دارد.
- موجودیت، ارتباطهایی با موجودیتهای دیگر دارد.

موجودیت مستقل و موجودیت وابسته:

■ **موجودیت مستقل (قوی)**، موجودیتی است که مستقل از هر موجودیت دیگر و به خودی خود، در یک محیط مشخص مطرح باشد.

■ **موجودیت وابسته (ضعیف)**، موجودیتی است که وجودش وابسته به یک نوع موجودیت دیگر است.

تعریف صفت:

صفت در واقع خصیصه یا ویژگی یک نوع موجودیت است و هر نوع موجودیت مجموعه‌ای از صفات دارد. هر صفت یک نام، یک نوع و یک معنای مشخص دارد.

مثال:

موجودیت: دانشجو، درس، استاد و ...

صفت: نام درس، شماره درس، تعداد واحد و ...

انواع صفت:

ساده یا مرکب

تک مقداری یا چند مقداری

شناسه یا ناشناسه

هیچ مقدار پذیر یا هیچ مقدار ناپذیر

ذخیره شده یا مشتق

انواع
صفت

صفت ساده, مرکب, تک مقداری و چند مقداری:

- **صفت ساده** صفتی است که مقدارش از لحاظ معنایی تجزیه‌نشدنی و ساده باشد. (درس)
- **صفت مرکب** صفتی است که از چند صفت ساده تشکیل شده است. (آدرس: نام استان و شهر و ...)
- **صفت تک‌مقداری** صفتی است که برای یک نمونه از یک نوع موجودیت حداکثر یک مقدار از دامنه مقادیر را می‌گیرد. (شماره درس)
- **صفت چندمقداری** صفتی است که بیش از یک مقدار از دامنه مقادیر می‌گیرد. (مدرک دانشگاهی: لیسانس و دکترا)

صفت شناسه یا ناشناسه:

- صفت شناسه موجودیت که گاه به آن کلید هم گفته می‌شود، دو ویژگی دارد:

۱- یکتایی مقدار دارد. ۲- حتی الامکان طول مقادیرش کوتاه است.

(مانند: شماره ملی , شماره کارگزینی)

هر صفتی که شناسه نباشد **ناشناسه** است.

صفت هیچ مقدارپذیر یا هیچ مقدارناپذیر:

- اگر صفت **هیچ مقدارپذیر** باشد می تواند مقدار نداشته باشد. (شماره تلفن)
- صفت **هیچ مقدارناپذیر** صفتی است که باید مقدار داشته باشد. (صفت شناسه)

صفت ذخیره شده (واقعی) و صفت مشتق:

■ صفت **ذخیره شده** صفتی است که مقادیرش در پایگاه داده‌ها ذخیره شده باشد.
(نام خانوادگی)

■ صفت **مشتق** صفتی است که مقادیرش در پایگاه داده‌ها ذخیره نشده باشد، بلکه حاصل یک پردازش روی فقره‌هایی از داده‌های ذخیره شده باشد. (معدل)

تعریف ارتباط:

ارتباط اندرکنش (تعامل) بین دو یا بیش از دو نوع موجودیت است و ماهیتاً نوعی بستگی بین انواع موجودیتهاست.

مثال:

دانشجو درس را انتخاب می کند.

دانشجو درس را حذف می کند.

جلسه سوم

نمودار ER

آنچه در این جلسه می خوانید:

- ۱- نمودار ER
- ۲- نمادهای نمودار ER
- ۳- وضع مشارکت در ارتباط
- ۴- درجه ارتباط
- ۵- چندی ارتباط

نمودار ER:

نمودار ER نموداری است که سه مفهوم اساسی مدل ER، یعنی نوع موجودیت، صفت و ارتباط نمایش داده می‌شوند.

نمادهای نمودار ER:

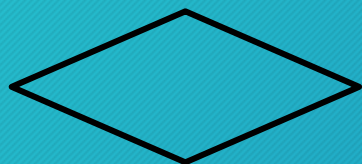
موجودیت



موجودیت ضعیف



ارتباط



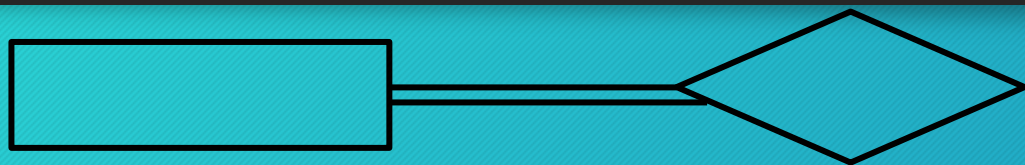
ارتباط با موجودیت ضعیف



مشارکت موجودیت در ارتباط



نمادهای نمودار ER:



مشارکت الزامی



صفت



صفت شناسه اول



صفت شناسه دوم

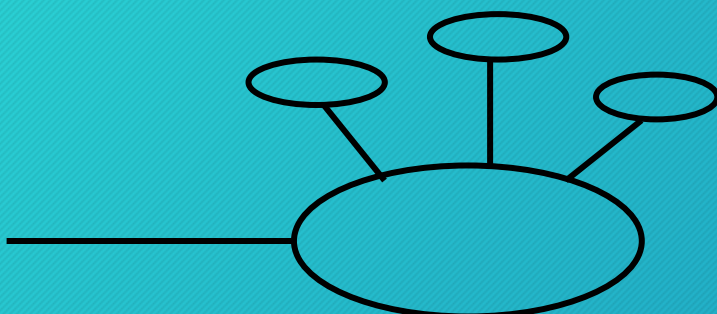


صفت شناسه مرکب



صفت چندمقداری

نمادهای نمودار ER:



صفت مرکب



صفت مشتق



چندى ارتباط
1 به N
1 به 1
N به M



ارتباط "گونه‌ای است از"

وضع مشارکت در ارتباط:

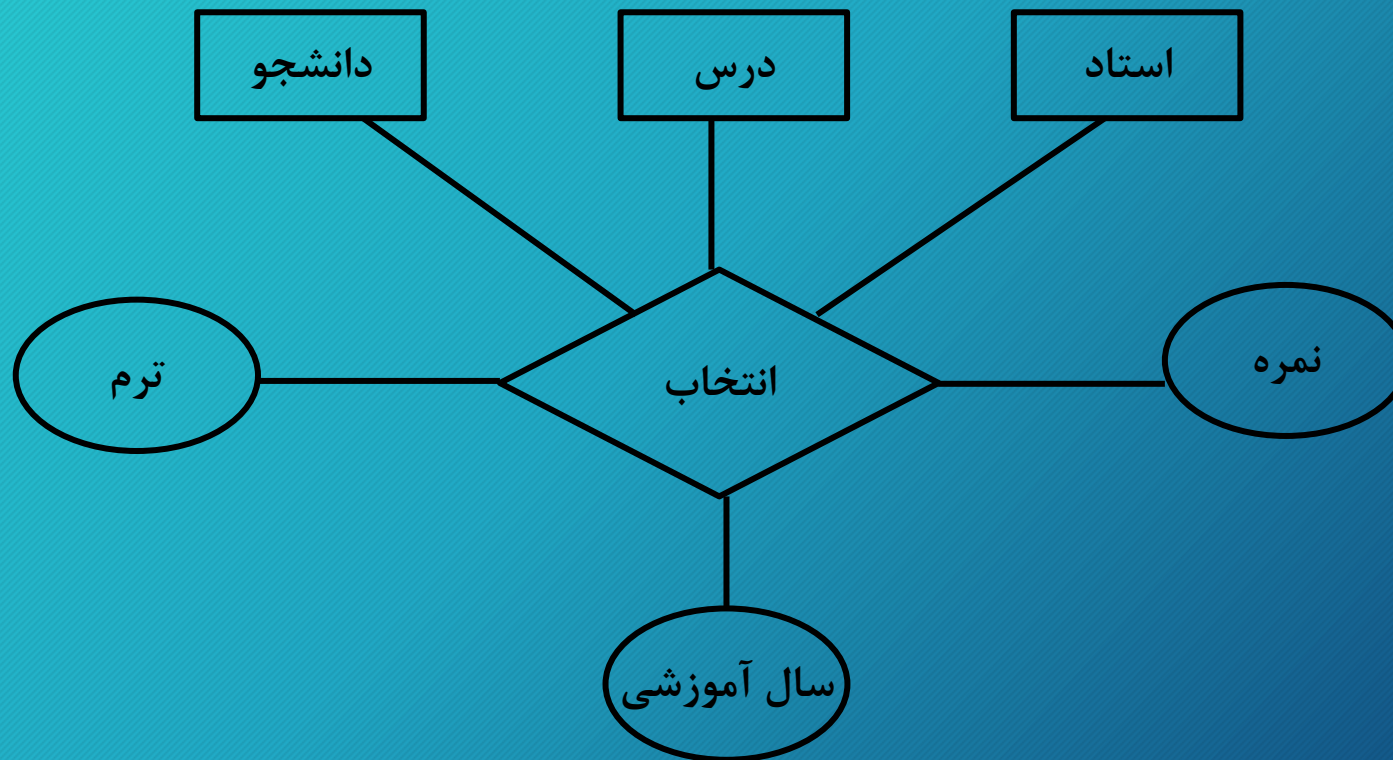
مشارکت یک نوع موجودیت در یک نوع ارتباط را الزامی گویند، اگر تمام نمونه‌های آن نوع موجودیت در آن نوع ارتباط شرکت کنند. در غیر این صورت مشارکت غیرالزامی است.



نمایش مشارکت
الزامی

درجه ارتباط:

تعداد شرکت کنندگان در یک ارتباط را درجه آن ارتباط می گویند.



ارتباط
بین سه
موجودیت

چندی ارتباط:

چندی یا ماهیت نوع ارتباط عبارتست از چگونگی تناظر بین دو مجموعه نمونه‌های آن دو نوع موجودیت.

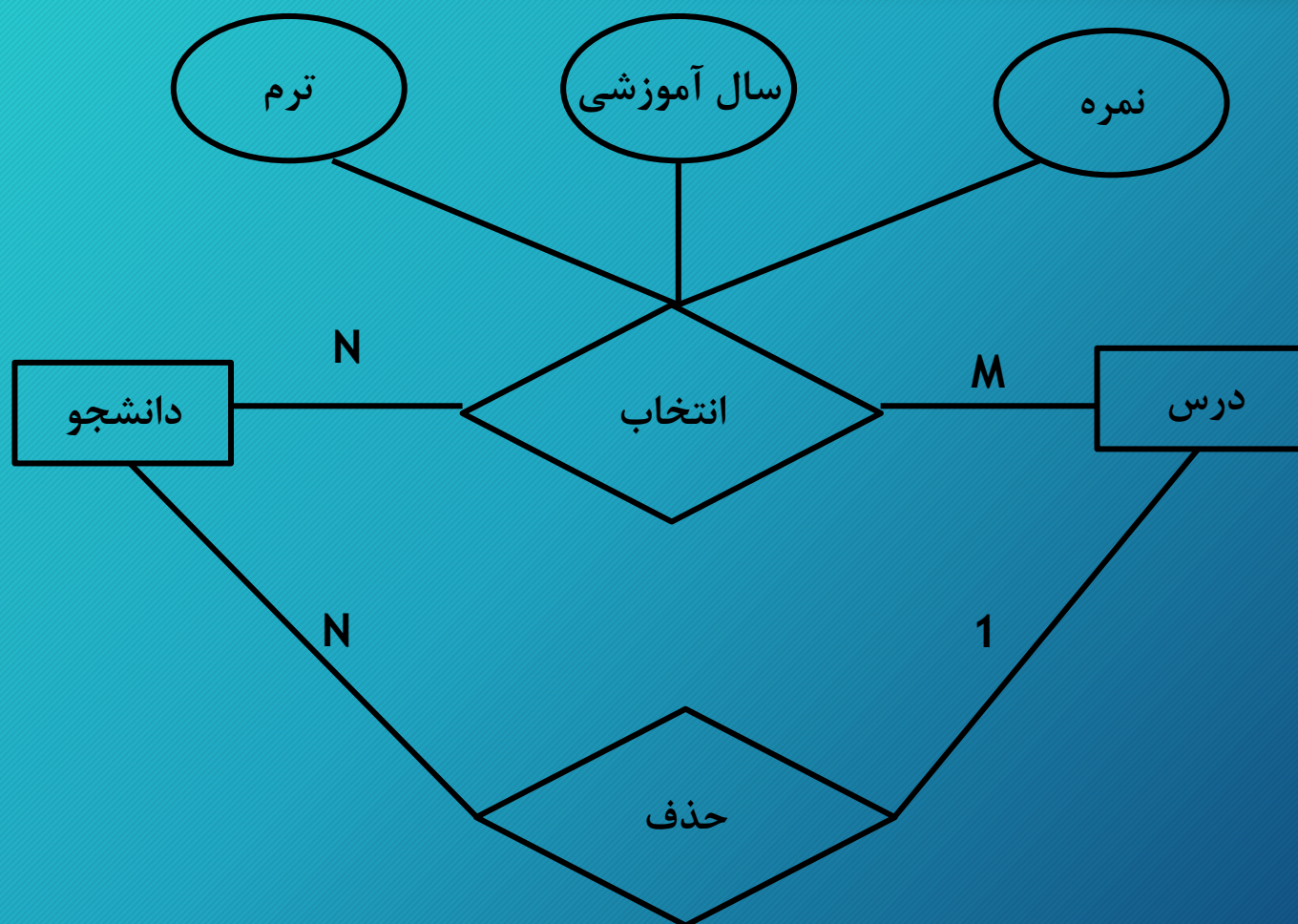
یک به یک 1:1

یک به چند 1:N

چند به چند N:M

انواع چندی ارتباط:

مثال چندی ارتباط:



محدودیت‌های روش ER

آنچه در این جلسه می خوانید:

- ۱- مشکلات روش ER (سه نوع دام)
- ۲- تجزیه و ترکیب
- ۳- تخصیص و تعمیم
- ۴- تجمع
- ۵- وراثت صفت
- ۶- دسته بندی
- ۷- مراحل مدل سازی معنایی داده ها
- ۸- روش مدل سازی UML
- ۹- نمادها
- ۱۰- خصوصیات کلی روش مدل سازی معنایی داده ها

مشکلات روش ER:

۱- دام حلقه ای:

هنگامی ایجاد می شود که با داشتن مثلا سه ارتباط دو موجودیتی، وجود یک ارتباط سه موجودیتی را نتیجه بگیریم در وضعی که این استنتاج درست نباشد.

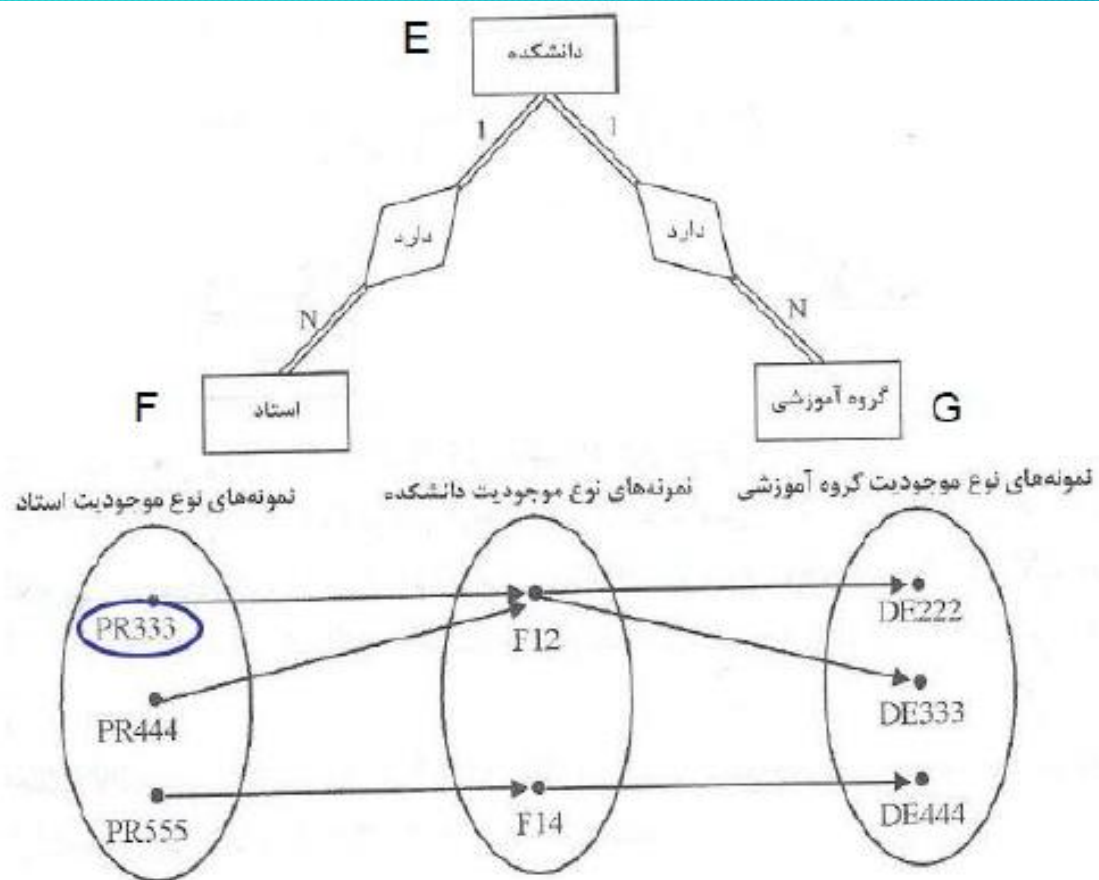
- دانشجوي x درس y را انتخاب کرده است.
- درس y توسط استاد z درس داده می شود.
- استاد z به دانشجوي x درس می دهد.
- نتیجه نمی توان گرفت که:
دانشجوي x درس y را با استاد z گرفته است.

مشکلات روش ER:

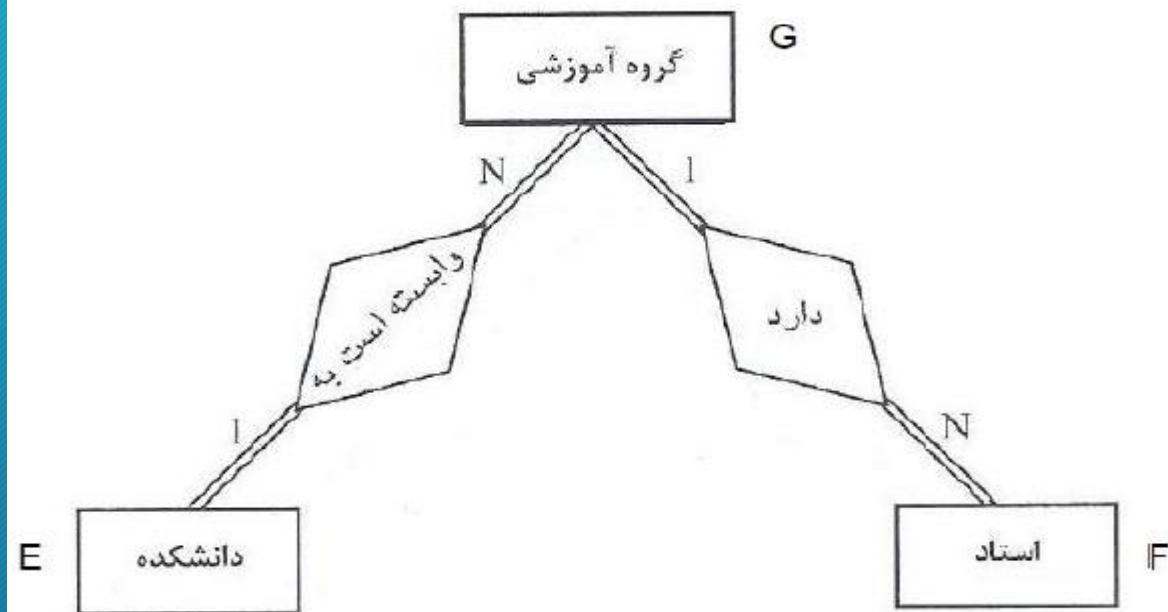
۲- دام چند شاخه (چتری):

این نوع دام وقتی ایجاد می شود که بین یک نوع موجودیت E و هریک از دیگر انواع موجودیت F، G و ... ارتباط 1:N با مشارکت الزامی وجود داشته باشد، ولی ارتباط بین مثلاً F و G، در مدلسازی دیده نشده باشد. در این صورت نمی توان با داشتن ارتباط های دوگانی بین E و F و بین E، G و ... اطلاع دو موجودیتی در مورد ارتباط بین نمونه های دو نوع موجودیت F و G را به دست آورد. اگر مدل ساز تصور کند که چنین اطلاعی را می تواند به دست آورد، دچار دام چندشاخه می شود.

مشکلات روش ER: (دام چند شاخه)



رفع دام چند شاخه



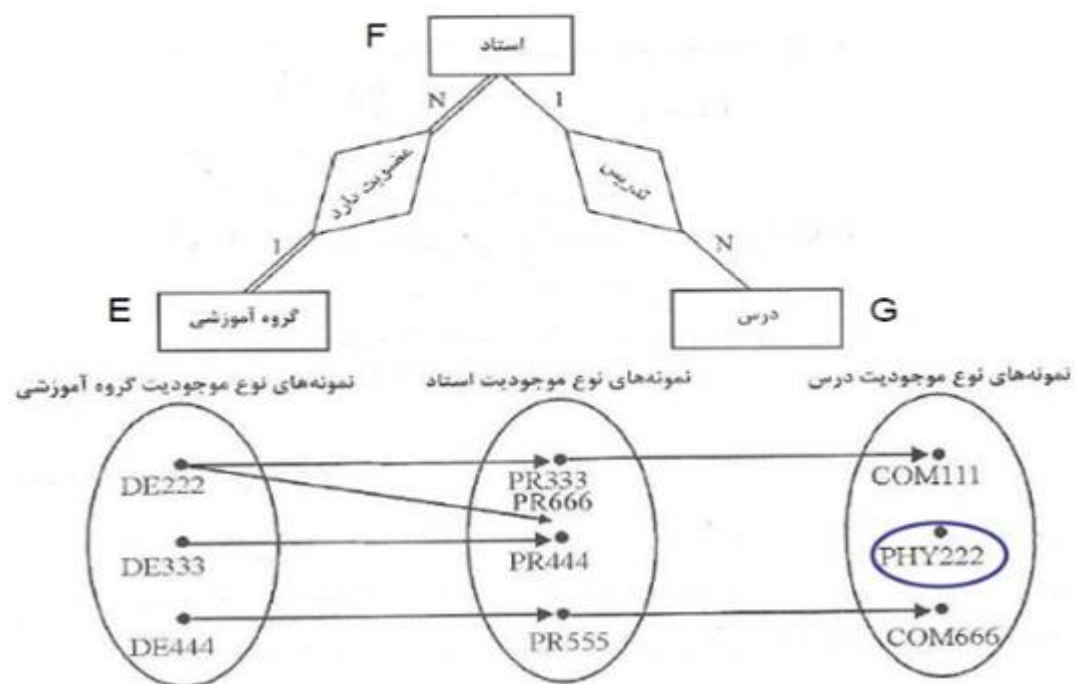
مشکلات روش ER:

۳- دام شکاف (گسل):

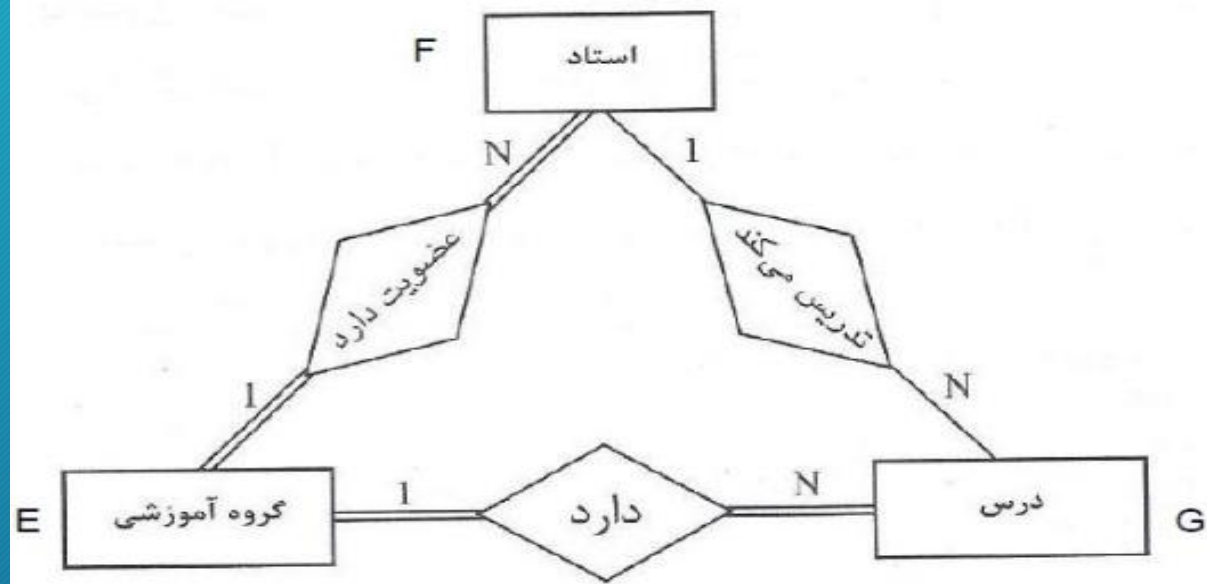
هنگامی ایجاد می‌شود که بین دو نوع موجودیت E و F، یک ارتباط باچندی 1:N و مشارکت الزامی وجود داشته باشد، ولی F خود با نوع موجودیت G، ارتباط 1:N با مشارکت غیرالزامی داشته باشد. به دلیل غیر الزامی بودن ارتباط بین F و G، نمی‌توان همه اطلاعات دو موجودیتی در مورد ارتباط بین نمونه‌های دو نوع موجودیت E و G را بدست آورد.

مشکلات روش ER: (دام شکاف)

دام شکاف



رفع دام گسل



محدودیت‌های روش ER که در روش EER برطرف گردیدند:

- ۱- تجزیه
- ۲- ترکیب
- ۳- تخصیص
- ۴- تعمیم
- ۵- تجمع
- ۶- وراثت صفت

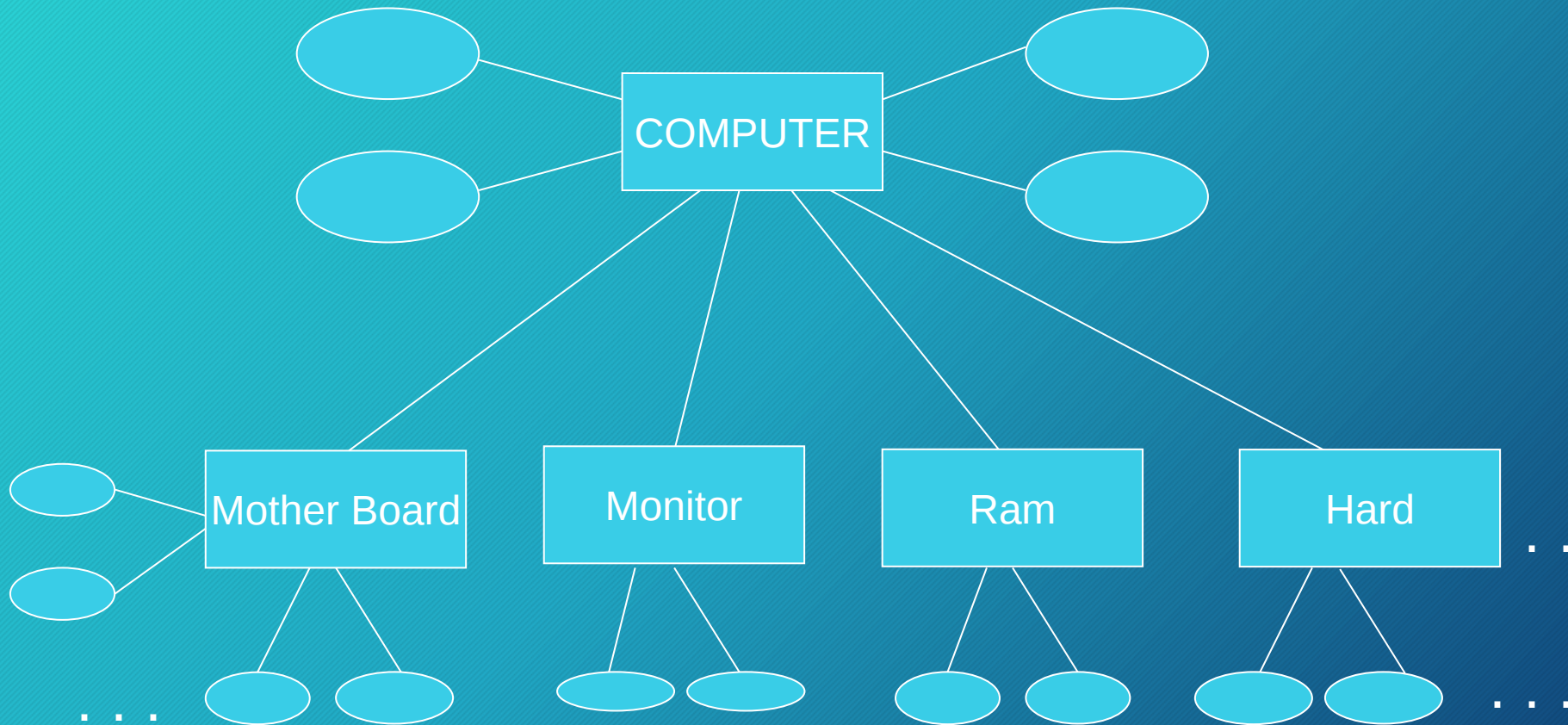
تجزیه:

تجزیه یا جداسازی یعنی یک شیء کل را به اجزاء تشکیل دهنده آن تقسیم کنیم. شیء کل صفات، ساختار و رفتار خود را دارد و هریک از اجزاء نیز صفات، ساختار و رفتار خاص خود را دارند. شیء کل شامل اجزاء خود است و بین شیء کل و اجزایش، ارتباط شمول وجود دارد. به این نوع ارتباط در EER، ارتباط "جزئی است از ..." گفته می شود.

ترکیب:

ترکیب، عکس عمل تجزیه است و در این عمل، با داشتن $E_i (i=1, 2, \dots)$ یک نوع موجودیت E را بازشناسی می‌کنیم به نحوی که E_i ها اجزاء تشکیل دهنده آن باشند.

مثال تجزیه و ترکیب:



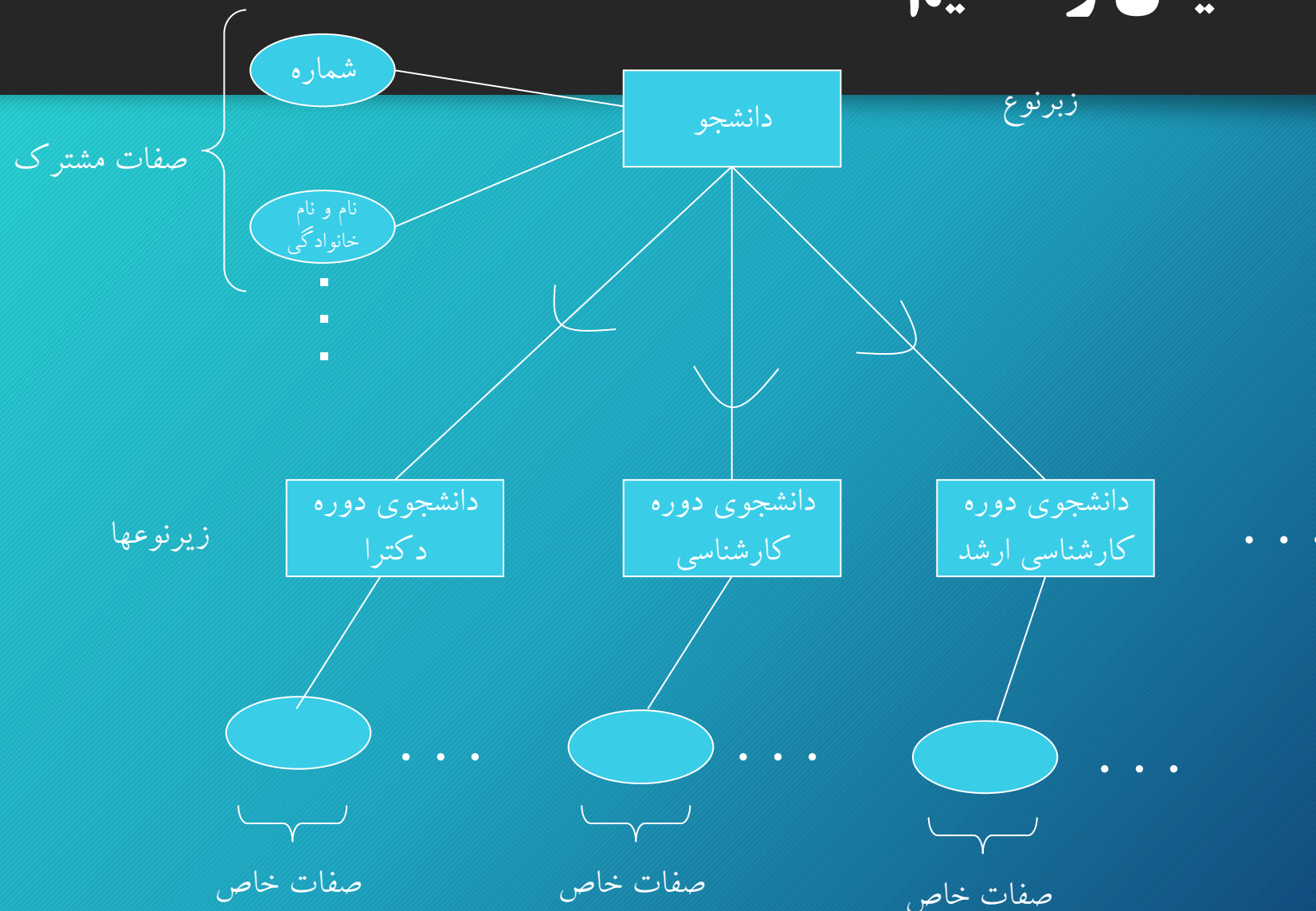
تخصیص:

تخصیص عبارتست از مشخص کردن گونه‌های خاص یک شیء براساس یک یا چند ضابطه مشخص، مثلاً اگر شیء موجود زنده را در نظر بگیریم، سه گونه خاص آن عبارتند از: انسان، حیوان و نبات. در روش EER هر یک نوع موجودیت می‌تواند خود زیرنوع موجودیتهایی داشته باشد. بین هر زیرنوع و زیرنوع ارتباط "گونه‌ای است از ..." وجود دارد.

تعمیم:

تعمیم، عکس عمل تخصیص است، به این معنا که با داشتن زیرنوعهای خاص، صفات مشترک بین آنها را در یک مجموعه صفات برای یک زیرنوع موجودیت در نظر می‌گیریم.

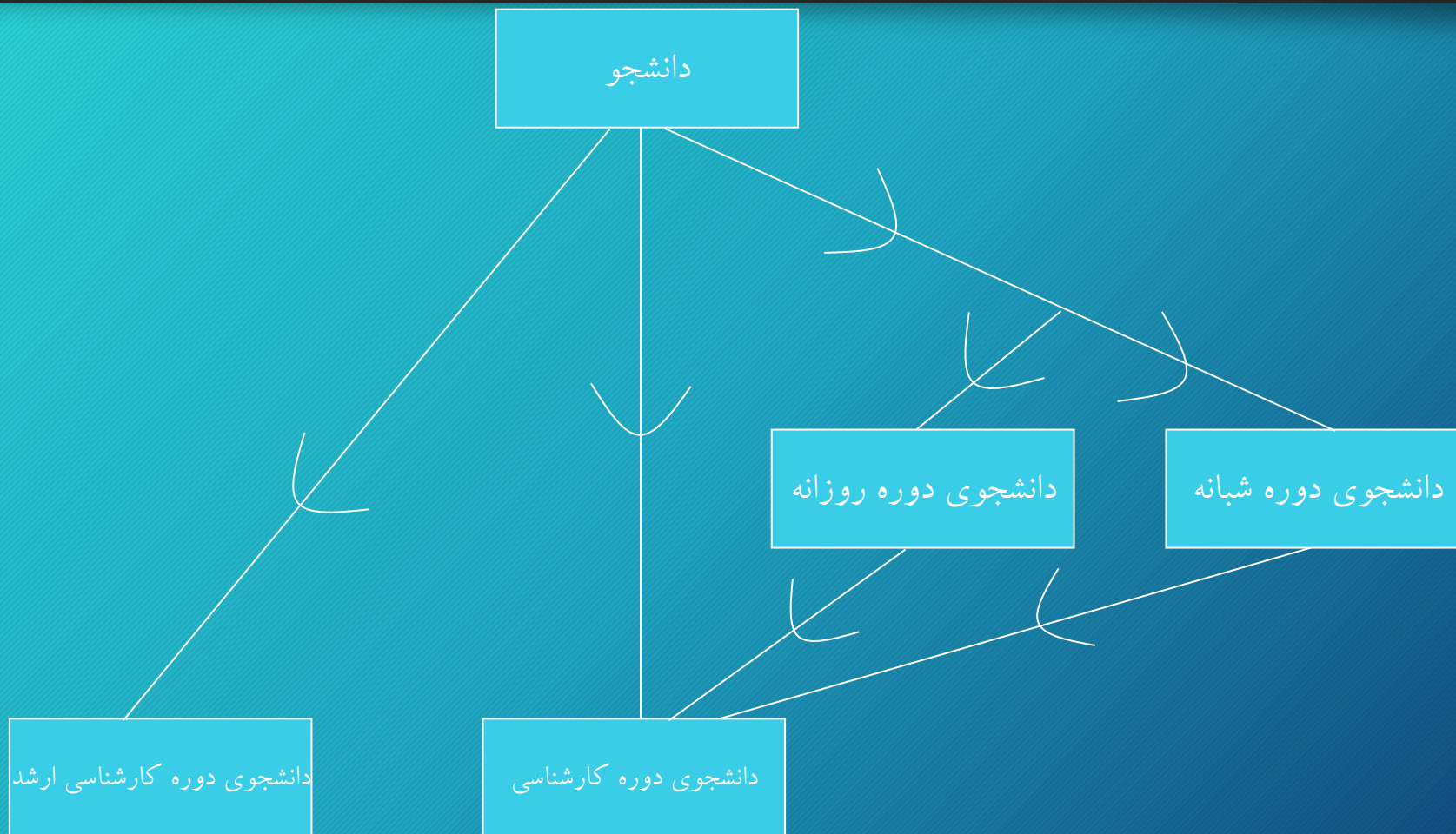
مثال تخصیص و تعمیم:



وراثت چندگانه:

یک زیرنوع موجودیت، می تواند در عین حال زیرنوع یک زیرنوع موجودیت دیگر هم باشد. با این ترتیب می توان مفهوم وراثت چندگانه را در روش EER نمایش داد.

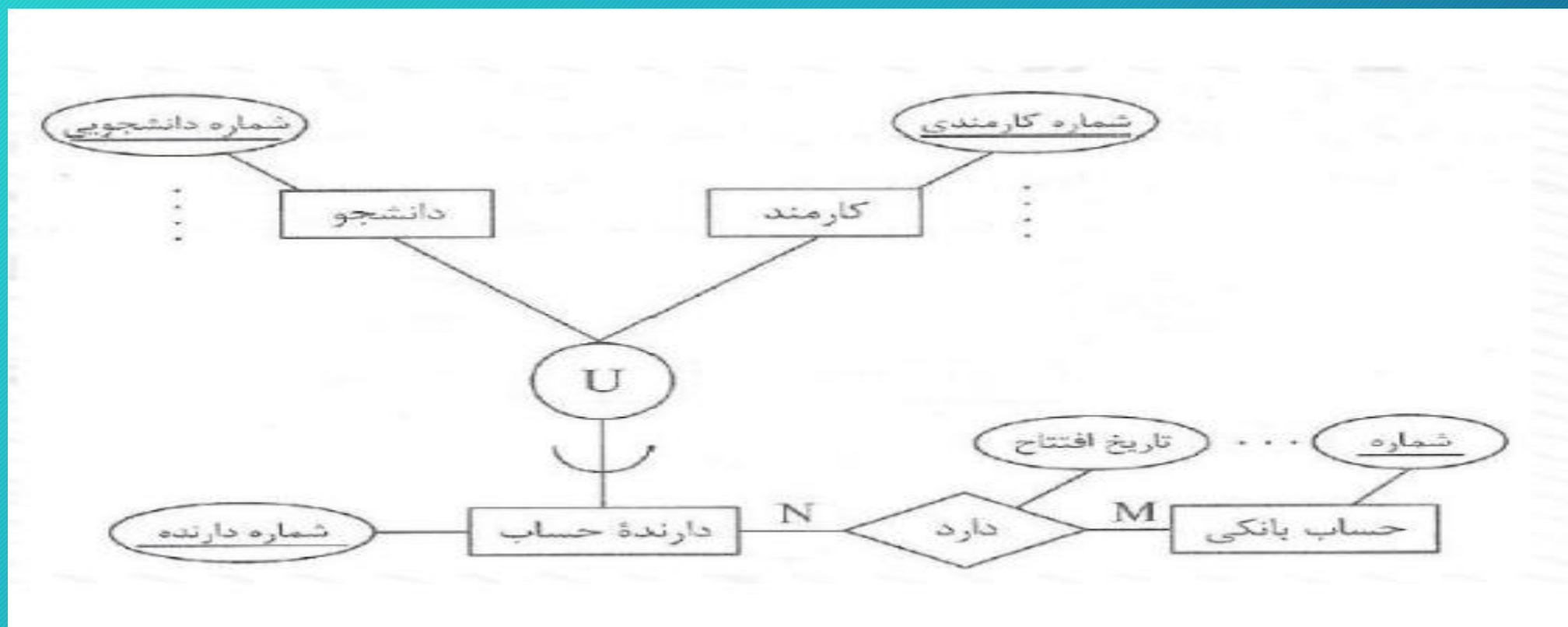
مثال وراثت چندگانه:



دسته بندی:

یک زیرنوع می تواند زیرنوع بیش از یک زیرنوع باشد. ممکن است زیرنوعهای این زیرنوع، از یک نوع نباشند. به این زیرنوع اصطلاحاً دسته (طبقه) گویند. برای نمایش دسته، از نماد U استفاده می شود.

مثال دسته بندی:

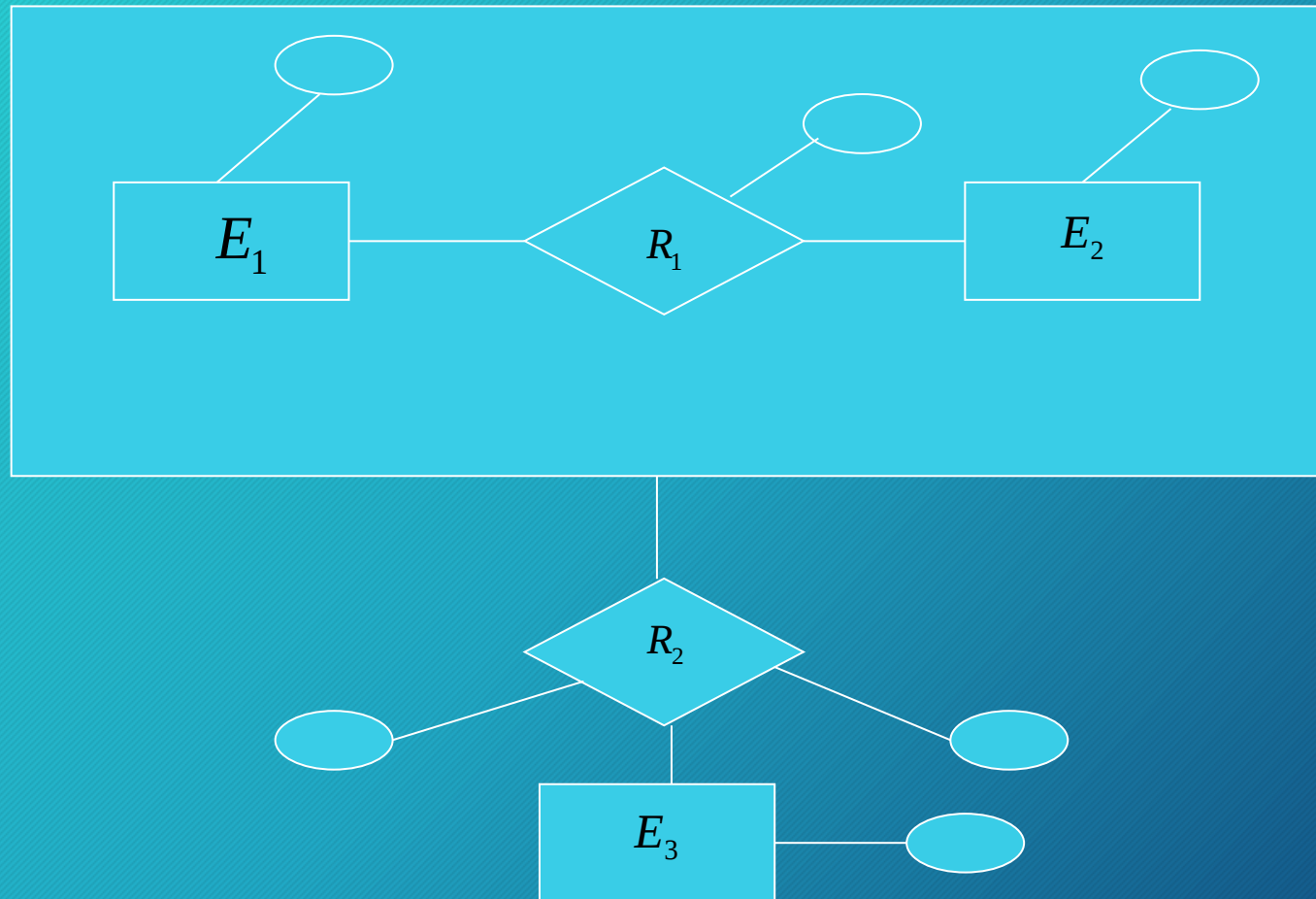


تجمع:

تجمع عبارتست از ساختن یک نوع موجودیت جدید با دیدن دو یا بیش از دو نوع موجودیت، که خود باهم در یک ارتباط شرکت دارند، به صورت یک نوع موجودیت واحد. در واقع مجموعه‌ای از موجودیتهای مرتبط را باهم مجتمع کرده و به عنوان یک نوع موجودیت واحد، در نظر می‌گیریم و این نوع موجودیت واحد خود می‌تواند با نوع موجودیت دیگری ارتباط داشته باشد.

مثال تجمع:

E_k



مراحل مدل سازی معنایی داده ها:

- ۱- مطالعه، تحلیل و شناخت محیط
- ۲- برآورد خواسته ها و نیازهای اطلاعاتی و پردازشی همه کاربران و تشخیص محدودیتهای معنایی
- ۳- بازشناسی انواع موجودیتهای مطرح و تعیین وضع هریک
- ۴- تعیین مجموعه صفات هر نوع موجودیت
- ۵- بازشناسی انواع ارتباطات بین انواع موجودیتهای تشخیص نوع مشارکت و چندی ارتباط
- ۶- رسم نمودار ER
- ۷- فهرست کردن پرسشهایی که پاسخ آنها از نمودار ER بدست می آید.
- ۸- واریسی مدلسازی انجام شده تا اطمینان حاصل شود که مدلسازی پاسخگوی نیاز کاربران است.

روش مدلسازی UML:

در این روش از چند نمودار برای نمایش مدلسازی و طراحی نرم افزار استفاده می شود.

مفاهیم اصلی در این مدلسازی:

- رده
- صفت
- بستگی

تناظر بین مفاهیم UML و مفاهیم EER

مفهوم در EER	مفهوم در UML
نوع موجودیت	رده
نمونه موجودیت	شیء
صفت	صفت
ارتباط	بستگی
نمونه ارتباط	پیوند
ارتباط بازگشتی	بستگی انعکاسی
نوع موجودیت ضعیف	بستگی مقید
صفت مرکب	میدان ساختمند
صفت ارتباط	صفت پیوند
چندی ارتباط	چندی بستگی

نمادها:

رده با یک مربع یا مستطیل نشان داده می شود و به شکل زیر به سه قسمت تقسیم می شود:

نام رده
نام صفات
پردازشها

نمادها:

بستگی بین دو رده، به صورت یک خط متصل کننده دو رده نمایش داده می شود و نام ارتباط روی خط نوشته می شود.

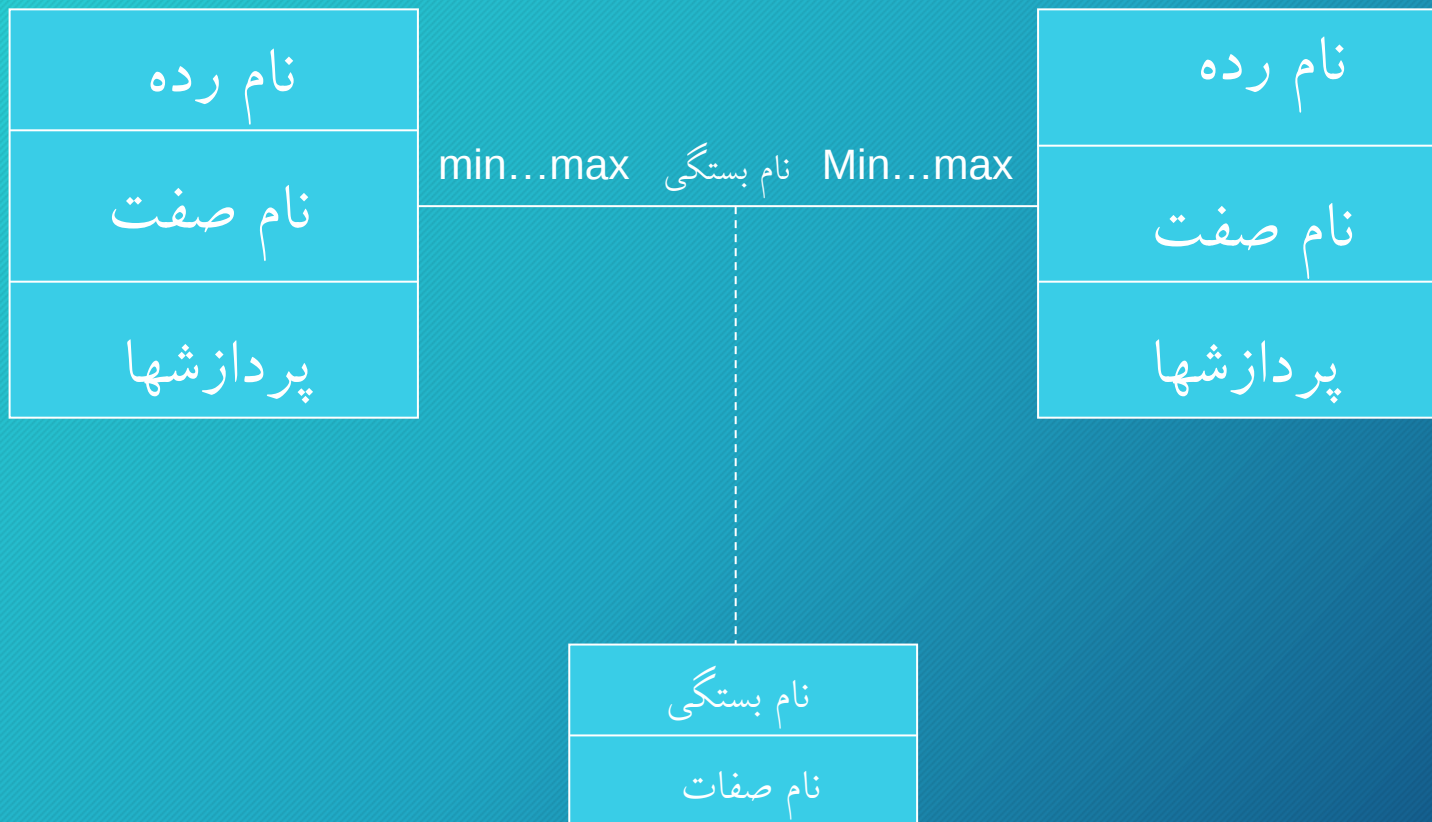
تجمع به صورت زیر نشان داده می شود:



نکته:

- چندی بستگی به صورت $\min...max$ نوشته می شود. اگر به جای max علامت ستاره باشد، به این معنا است که مقدار max محدودیت ندارد.
- صفت چندمقداری به صورت یک رده جداگانه نشان داده می شود، ولی فاقد قسمت مربوط به پردازش است.
- نام میدان بعد از نام صفت نوشته می شود و بین این دو نام علامت : گذاشته می شود.
- صفت پیوندی در یک مربع (مستطیل) نوشته می شود و این مربع با خط چین به خط نشان دهنده بستگی متصل می شود. نام پیوند و نام صفات پیوند در دو قسمت از این مربع گذاشته می شود.

نمایش صفت پیوند در UML:



خصوصیات کلی روش مدل سازی معنایی داده ها:

- گویایی
- صوری بودن
- سادگی مفاهیم
- قابلیت نمایش نموداری
- ایجاز
- جامع بودن مفاهیم
- قابلیت نمایش ساختار، حالت و رفتار نوع موجودیت
- گسترش پذیری

جلسه پنجم

آشنایی با ساختار داده ها

آنچه در این جلسه می خوانید:

- ۱- پایگاه داده‌ها در محیط انتزاعی
- ۲- سطوح محیط انتزاعی
- ۳- گونه‌های موجود ساختار داده‌ای
- ۴- مفهوم ساختار داده‌ای در سطوح مختلف پایگاه داده‌ها
- ۵- ساختار داده‌ای رابطه‌ای
- ۶- شمای پایگاه جدولی
- ۷- عملیات در پایگاه جدولی
- ۸- عملگرهای جبر رابطه‌ای
- ۹- برخی ویژگیهای ساختار داده‌ای جدولی

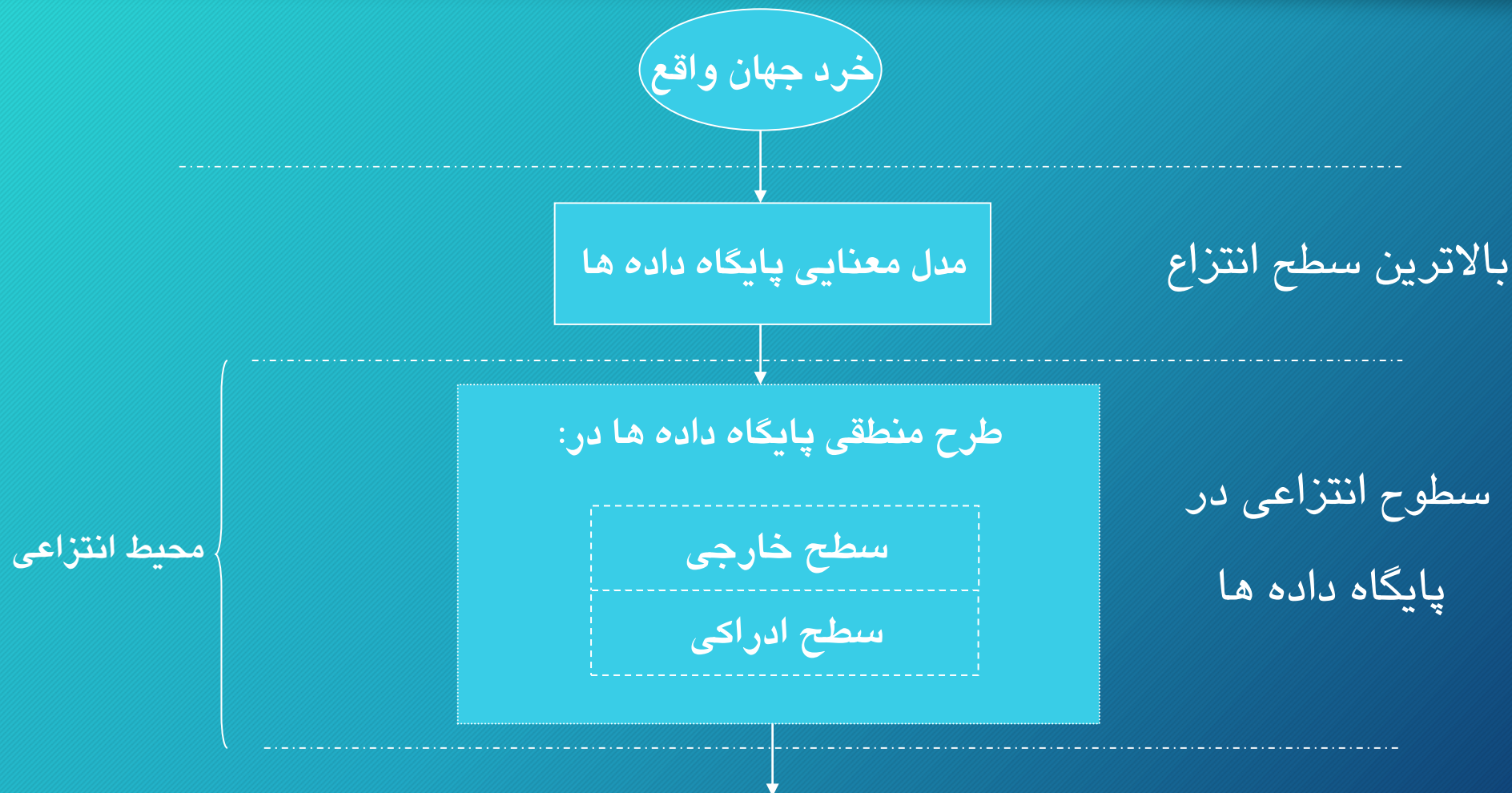
پایگاه داده‌ها در محیط انتزاعی:

در جلسات قبل دیدیم که مدل سازی معنایی داده ها, نمایی از پایگاه داده ها در یک محیط انتزاعی است که این محیط, محیطی است فراتر از محیط فایلینگ منطقی و فایلینگ فیزیکی است و مفاهیمی که در این محیط مطرح می‌شوند باید از جنبه‌های فایلینگ پایگاه مستقل و ماهیتاً انتزاعی باشند.

مفهوم محیط انتزاعی می‌تواند سطوحی داشته باشد:

- ۱- بالاترین سطح انتزاع (مدل معنایی پایگاه داده ها)
- ۲- سطح انتزاعی پایین تر (طراحی منطقی پایگاه داده ها)

سطوح محیط انتزاعی:



طراحی منطقی پایگاه داده ها:

همانطور که در مدلسازی داده ها نیاز به امکاناتی برای نمایش واقعیات بود (نمودارها) , برای طراحی منطقی پایگاه داده ها هم امکانات خاصی لازم است مانند یک مدل داده که شامل ساختار داده ای باشد تا طراحی منطقی پایگاه داده ها با استفاده از مفاهیم اساسی یک مدل داده و در چارچوب ساختار داده آن مدل انجام شود.

ساختار داده:

ساختار داده امکانی است برای نمایش نوع موجودیت ها و ارتباطهای بین آنها در سطح طراحی منطقی پایگاه داده ها.

نکته:

هر ساختار داده حداقل یک عنصر ساختاری اساسی دارد.
عنصر اساسی ساختاری عنصری است که به کمک آن نوع موجودیت یا نوع ارتباط یا هر دو نمایش داده می شوند.

انواع ساختارهای داده:

- ۱- ساختار داده‌ای رابطه‌ای (جدولی)
- ۲- ساختار داده‌ای سلسله مراتبی
- ۳- ساختار داده‌ای شبکه‌ای
- ۴- ساختار داده‌ای جعبه‌ای
- ۵- ساختار داده‌ای هایپرگرافی
- ۶- ساختار داده‌ای لیستهای وارون

آشنایی با ساختار دادهٔ رابطه ای (جدولی):

پایگاه داده ای جدولی مجموعه ای است از تعدادی جدول و هر جدول یک نام و حداقل یک ستون با نام و نوع داده مشخص دارد که جدول همان عنصر ساختاری اساسی است.

مفاهیم اصلی ساختار جدولی عبارتند از:

۱- جدول

۲- سطر

۳- ستون

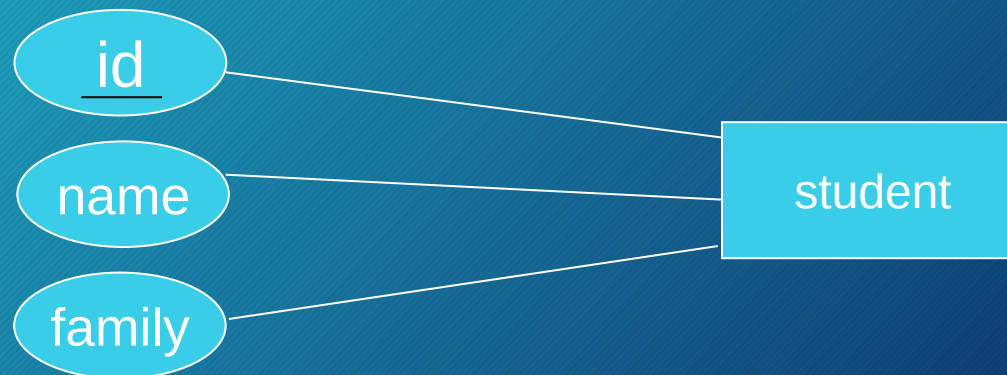
شمای پایگاه داده جدولی:

شمای پایگاه داده‌ها عبارتست از تعریف (توصیف) ساختهای منطقی طراحی شده و نوعی برنامه است شامل دستورات "تعریف داده‌ها" و "کنترل داده‌ها"، که "عملیات در داده‌ها" در آن وجود ندارد.

مثال:

Create table student
(id char(8) not null,
name char(20),
Family char(20))
Primary key id;

<u>id</u>	name	family
...	...	...



عملیات در پایگاه جدولی:

۱- درج (INSERT)

۲- حذف (DELETE)

۳- بهنگام سازی (UPDATE)

۴- بازیابی (SELECT)

Insert

Into student

Values ('9625320','ali','Ahmadi');

id	name	family
9625320	ali	Ahmadi

حذف:

Delete

From student

Where id='9625320';

id	name	family

بهنگام سازی:

Update student

Set id='9825320'

Where id='9625320';

id	name	family
9825320	ali	Ahmadi

بازیابی:

```
Select *  
From student  
Where name='ali';
```

- عملگر گزینش یا تحدید (restrict):
این عملگر تعدادی سطر از یک جدول را بازیابی می کند.
- عملگر پرتو (project):
این عملگر تعدادی ستون از یک جدول را بازیابی می کند.
- عملگر پیوند طبیعی (natural join):
این عملگر دو جدول را به هم پیوند می زند.

برخی ویژگیهای ساختار داده‌ای جدولی:

- ۱- از نظر کاربر نمایش ساده‌ای دارد.
- ۲- محیطش منطقاً مسطح است.
- ۳- تنها عنصر ساختاری اساسی آن جدول است.
- ۴- همه چیز با فقره داده‌های ساده نمایش داده می‌شود.
- ۵- ارتباطات با چندیهیهای مختلف در آن قابل نمایش است.
- ۶- منطق بازیابی آن ساده است.
- ۷- ساختار منطقی دستور بازیابی آن ساده است.
- ۸- برای پرسشهای قرینه، رویه پاسخگوی قرینه دارد.
- ۹- مبنای تئوریک قوی دارد.
- ۱۰- تقارن دارد.
- ۱۱- افزونگی ناشی از ماهیت ساختار داده دارد.
- ۱۲- به شرط انجام طراحی درست، در عملیات ذخیره سازی کمترین آنومالی (وضع غیر عادی و نامطلوب) را دارد.
- ۱۳- قواعد جامعیت ذاتی ندارد.

جلسه ششم

آشنایی با ساختار داده ها

آنچه در این جلسه می خوانید:

- ۱- آشنایی با ساختار داده سلسله مراتبی
- ۲- عملیات در پایگاه داده سلسله مراتبی
- ۳- برخی ویژگی های ساختار داده سلسله مراتبی
- ۴- آشنایی با ساختار داده شبکه ای
- ۵- عملیات در ساختار داده شبکه ای
- ۶- برخی ویژگی های ساختار داده شبکه ای
- ۷- چرا از ساختار داده استفاده می شود؟

ساختار داده سلسله مراتبی:

پایگاه داده سلسله مراتبی مجموعه ای است منطقاً منظم از نمونه های متمایز یک یا بیش از یک سلسله مراتب.

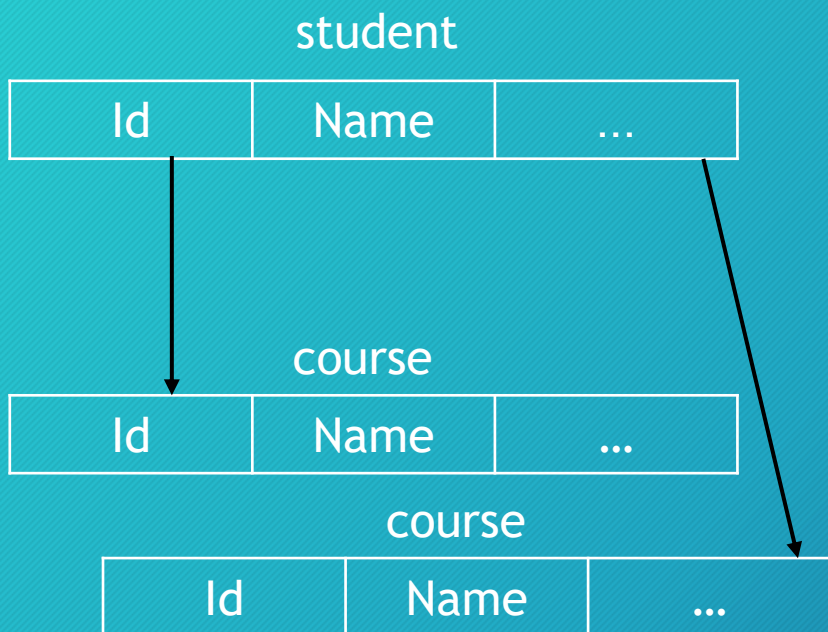
این ساختار قدیمی ترین ساختار داده برای طراحی منطقی پایگاه داده ها است.

سلسله مراتب در اساس نوعی درخت است.

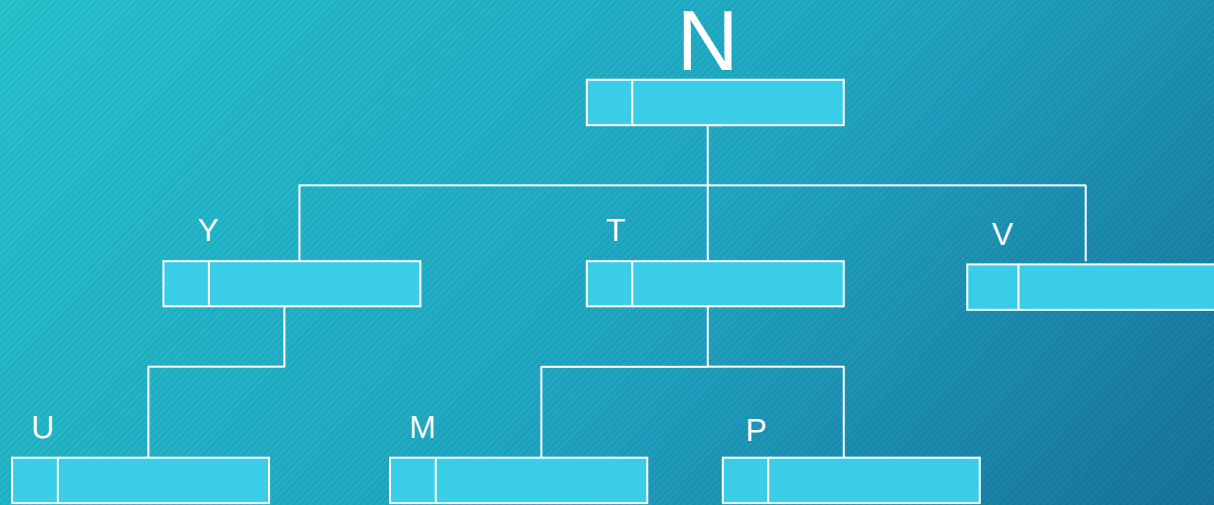
این ساختار دو عنصر ساختاری اساسی دارد:

۱- رکورد: برای نمایش موجودیت

۲- پیوند پدر - فرزندی: برای نمایش ارتباط دو موجودیت



مثال ساختار دادهٔ سلسله مراتبی:



عملیات در پایگاه دادهٔ سلسله مراتبی:

- ۱- بازیابی یک رکورد
- ۲- بازیابی رکورد بعدی
- ۳- بازیابی یک فرزند از یک پدر بلافاصله
- ۴- بازیابی فرزند بعدی از پدر
- ۵- درج یک فرزند زیر یک پدر
- ۶- حذف یک فرزند از یک پدر
- ۷- بهنگام سازی یک یا چند فرزند از یک پدر

برخی ویژگیهای ساختار دادهٔ سلسله مراتبی:

- ۱- سادگی نمایش ساختار جدولی را ندارد.
- ۲- مبنای ریاضی ندارد.
- ۳- دو عنصر ساختاری اساسی دارد.
- ۴- برای نمایش ارتباط یک به چند مناسب است.
- ۵- نمایش ارتباط چند به چند در آن دشوار است.
- ۶- ساخت منطقی رویه بازیابی آن به سادگی منطق رویه بازیابی در ساختار جدولی نیست.
- ۷- سادگی عملیات ذخیره‌سازی ساختار دادهٔ جدولی را ندارد.
- ۸- تقارن ساختار جدولی را ندارد.
- ۹- تعدادی قاعده جامعیت ذاتی دارد.
- ۱۰- در مواقعی که در ذخیره‌سازی نمونه‌های فرزندان، افزونگی پدید آید، پایگاه در معرض ناسازگاری قرار می‌گیرد.

ساختار داده شبکه‌ای:

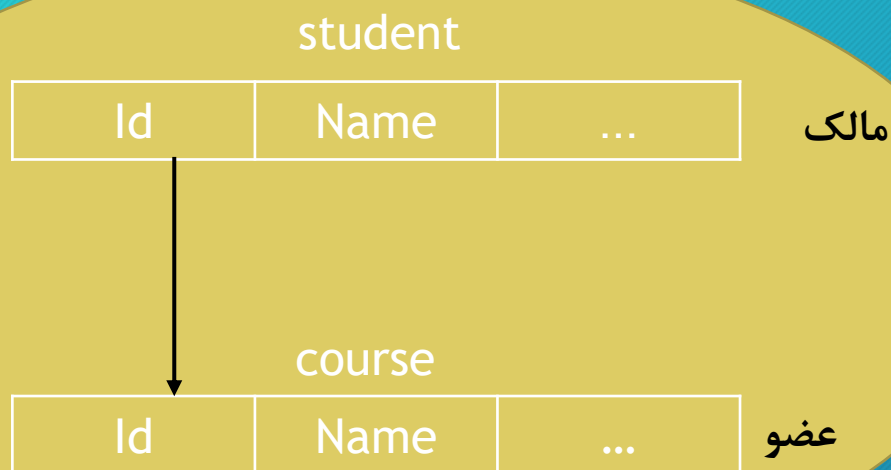
پایگاه داده شبکه ای مجموعه ای است منطقاً منظم از یک یا چند مجموعه. ساختار داده ای شبکه ای نوعی گراف جهت دار است و گره هایش به کمک یالهایی بهم وصل هستند. این ساختار را می توان گسترش یافته ساختار سلسله مراتبی دانست به این معنا که هر گره فرزند می تواند بیش از یک گره پدر داشته باشد.

این ساختار دو عنصر ساختاری اساسی دارد:

۱- رکورد: برای نمایش موجودیت

۲- مجموعه (کوداسیل): برای نمایش

ارتباط دو موجودیت



مجموعه یا کوداسیل انتخاب درس

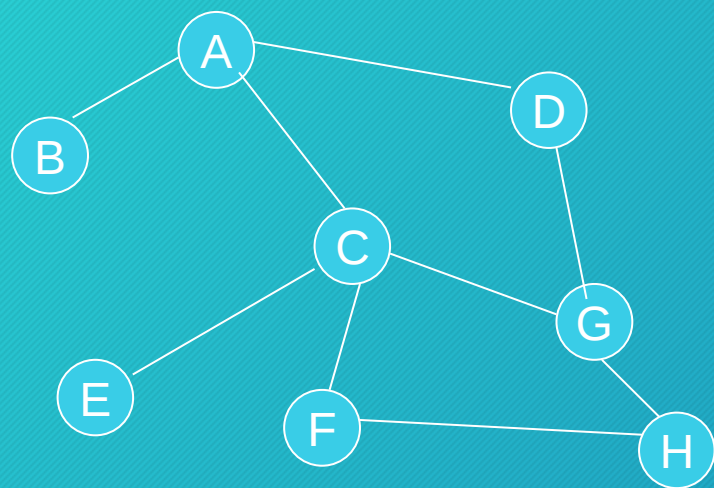
مثال ساختار داده شبکه ای:

در این ساختار مجموعه از سه قسمت تشکیل می شود:

۱- نام مجموعه

۲- رکورد مالک

۳- رکورد عضو



در این ساختار:

۱- مالک می تواند مالک مجموعه دیگر باشد.

۲- مالک می تواند عضو مجموعه دیگر باشد.

۳- عضو می تواند مالک مجموعه دیگر باشد.

۴- عضو می تواند عضو مجموعه دیگر باشد.

۵- مجموعه می تواند یک مالک و چندین عضو داشته باشد.

عملیات در پایگاه داده شبکه ای:

- ۱- بازیابی یک رکورد مالک
- ۲- بازیابی رکورد بعدی رکورد مالک
- ۳- بازیابی رکورد عضو با داشتن مالک
- ۴- بازیابی رکورد بعدی رکورد عضو با داشتن مالک
- ۵- درج یک رکورد مالک یا عضو
- ۶- حذف یک رکورد مالک یا عضو
- ۷- بهنگام سازی یک رکورد مالک یا عضو

برخی ویژگیهای ساختار دادهٔ شبکه ای:

- ۱- سادگی ظاهری ساختار داده‌ای جدولی را ندارد.
- ۲- مبنای ریاضی ندارد.
- ۳- دو عنصر ساختاری اساسی دارد.
- ۴- ماهیتاً برای نمایش ارتباطات "یک به چند" نیست، یعنی می‌توان ارتباط با چندیهای دیگر را هم نشان داد.
- ۵- ساخت منطقی دستور بازیابی آن پیچیده‌تر از ساختارهای دیگر است.
- ۶- مثل ساختار داده‌ای جدولی تقارن دارد.
- ۷- قواعد جامعیت ذاتی دارد.
- ۸- به علت حجم زیاد اشاره‌گرها، ایجاد یا اصلاح آنها می‌تواند سبب بروز فزونکاری در سیستم شود.
- ۹- بعضی آنومالیهای مدل سلسله مراتبی در عملیات ذخیره‌سازی را ندارد.

چرا ساختار داده:

- ۱- تأمین کننده محیط پایگاه داده ها است.
- ۲- مبنای طراحی دستورات تعریف داده , کنترل داده و دستورهای عملیات در داده ها است.
- ۳- مبنای طراحی سیستم مدیریت است.
- ۴- ضابطه ای است برای مقایسه سیستمهای مدیریت و نیز ارزیابی آنها.
- ۵- مبنایی است برای ایجاد و گسترش تکنیکهای طراحی پایگاه داده ها.
- ۶- مبنای تقسیم بندی و برنامه ریزی موضوعات آموزش دانش و تکنولوژی پایگاه داده ها است.
- ۷- مبنای پژوهش در دانش و تکنولوژی پایگاه داده ها است.

جلسه هفتم

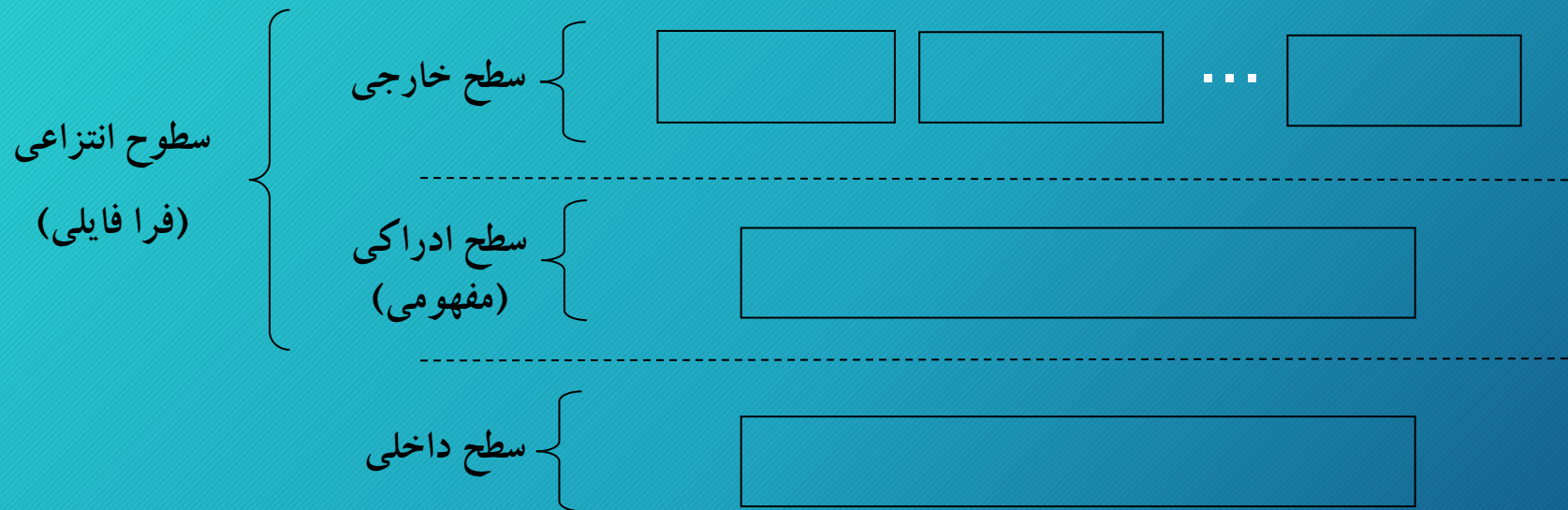
معماری پایگاه داده ها

آنچه در این جلسه می خوانید:

- ۱- معماری پیشنهادی ANSI
- ۲- دید ادراکی (مفهومی)
- ۳- دید خارجی
- ۴- دید داخلی
- ۵- زبان میزبان
- ۶- نقش افزایش تعداد زبانهای میزبان مورد پذیرش DBMS
- ۷- زبان دادهای فرعی
- ۸- دستورهای DSL برای سه سطح معماری پایگاه دادهها
- ۹- تقسیمبندی زبان دادهای فرعی از نظر نیاز به زبان میزبان
- ۱۰- ویژگیهای زبان دادهای فرعی

معماری پایگاه داده‌ها:

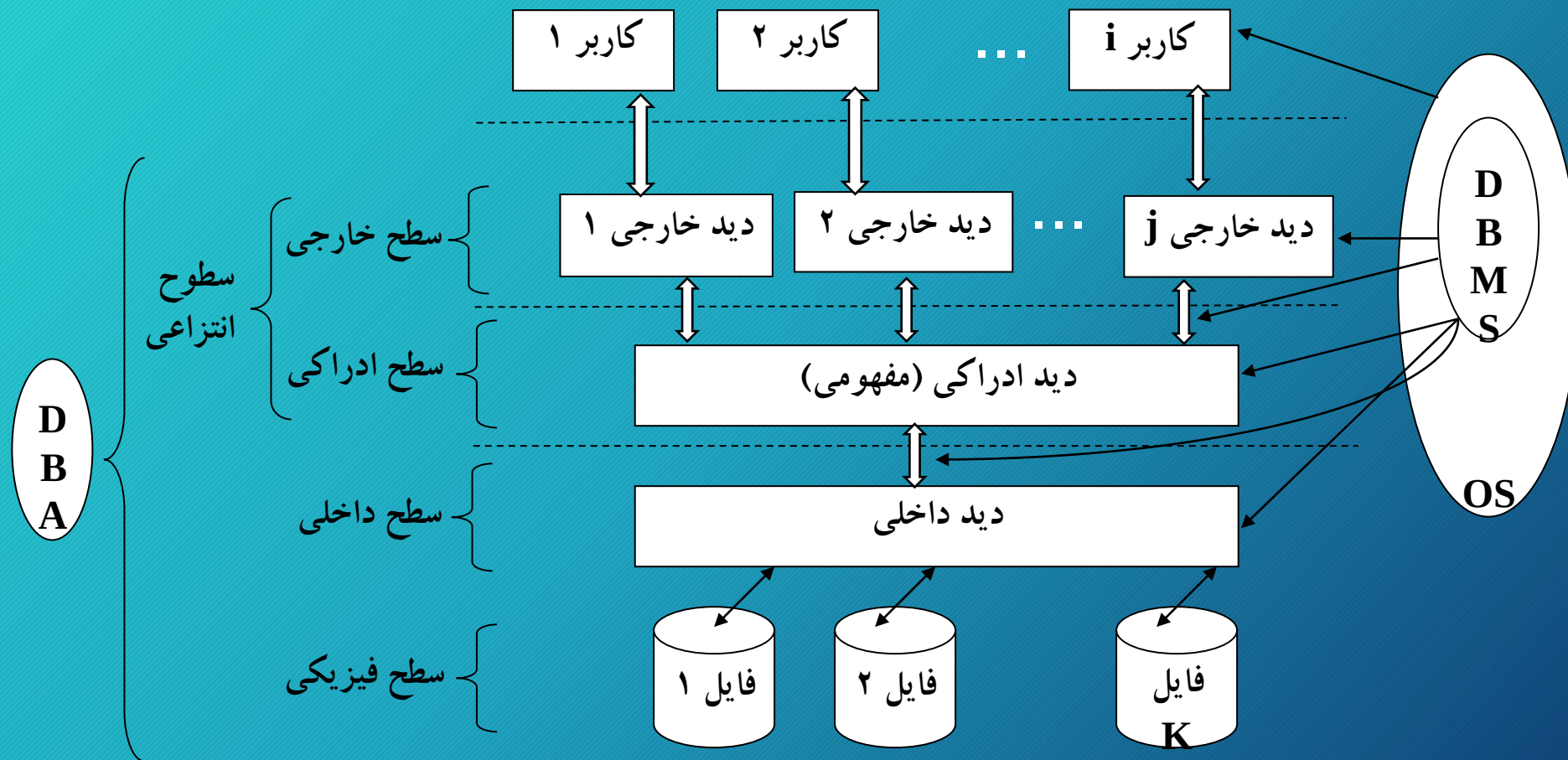
دید (نمای) ساده شده معماری سه سطحی



اجزای معماری پایگاه داده ها:

- ۱- کاربر
- ۲- زبان میزبان
- ۳- زبان داده ای فرعی
- ۴- دید خارجی
- ۵- دید ادراکی (مفهومی)
- ۶- دید داخلی
- ۷- فایل‌های فیزیکی
- ۸- سیستم مدیریت پایگاه داده ها
- ۹- مدیر پایگاه داده ها

اجزای معماری پایگاه داده ها:



دید ادراکی (مفهومی):

- ۱- دید طراح پایگاه داده‌ها نسبت به داده‌های ذخیره‌شده در پایگاه داده‌ها است.
- ۲- جامع (سرتاسری) است و در برگیرنده تمام نیازهای کاربران است.
- ۳- در یک محیط انتزاعی مطرح است بنابراین مبتنی بر یک ساختار داده مشخص است.
- ۴- با استفاده از عناصر ساختاری اساسی همان ساختار داده طراحی می‌شود.
- ۵- باید پس از طراحی توصیف شود.
- ۶- شمای ادراکی به سیستم داده می‌شود و در کاتالوگ سیستم نگهداری می‌شود.

مثال دید ادراکی (مفهومی):

student

Id	Name	Family

Create table student

Id char(20), Name char(20), Family char(20)

دید خارجی:

- ۱- دید کاربر خاص نسبت به داده‌های ذخیره شده در پایگاه داده است.
- ۲- جزئی است و نشان دهنده محدوده‌ای از پایگاه داده‌ها که به نیازهای داده‌ای یک کاربر خاص پاسخ می‌دهد.
- ۳- در سطح انتزاعی مطرح است و مبتنی بر یک ساختار داده مشخص.
- ۴- روی دید ادراکی طراحی و تعریف می‌شود.
- ۵- تعریف دید خارجی را شمای خارجی گفته و شمای خارجی نوعی برنامه است حاوی دستورات تعریف و کنترل داده‌ها در سطح خارجی که توسط کاربر این سطح نوشته می‌شود.
- ۶- به تعریف مجموعه دیدهای خارجی کاربر، سطح خارجی گفته می‌شود.
- ۷- هر کاربر می‌تواند تعدادی دید داشته باشد.
- ۸- چند کاربر می‌توانند در یک دید مشترک باشند.

دید داخلی:

- ۱- دید DBMS (سیستم مدیریت پایگاه داده ها) و طراح پایگاه داده ها است و در مرحله طراحی فیزیکی است و در سطحی پایین تر از سطح ادراکی، نسبت به کل داده های ذخیره شده است.
- ۲- در سطح فایلینگ منطقی مطرح است.
- ۳- مبتنی بر یک ساختار فایل است که با نظر و دخالت طراح پایگاه طراحی می شود و به طراحی فیزیکی معروف است.
- ۴- در سطح داخلی پایگاه داده ها، فایل های منطقی پایگاه داده ها تعریف می شود.
- ۵- شمای داخلی نوعی برنامه است که توسط خود DBMS تولید می شود و شرح فایلینگ منطقی پایگاه داده است.

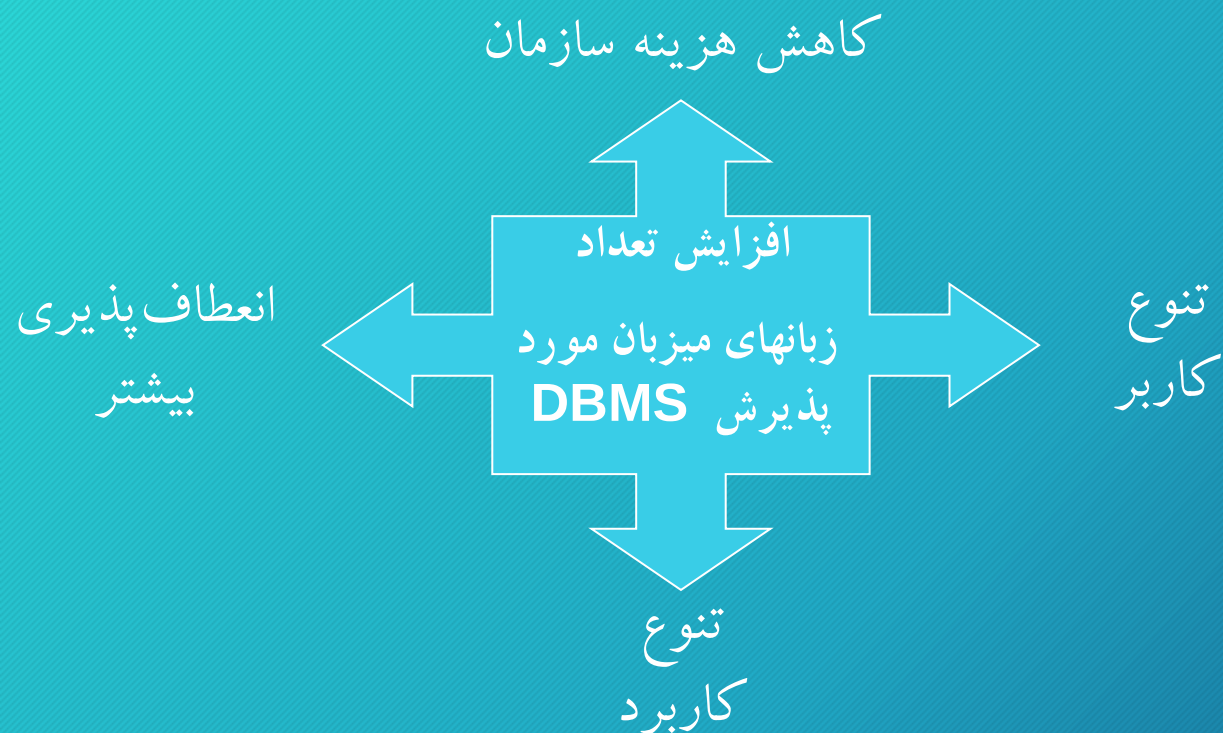
شرح اجزای دیگر

زبان میزبان:

یکی از زبانهای برنامه نویسی مانند کوبول , فرترن , پی ال وان , پاسکال , سی , اسمبلی و ... را زبان میزبان می گویند.

نکته:

هر چه تعداد زبانهای
مورد پذیرش سیستم
مدیریت بیشتر باشد
آن سیستم مطلوب تر است.



زبان داده‌ای فرعی:

این زبان از سه دسته دستور تشکیل شده است:

- ۱- دستورات تعریف داده‌ها (Data Definition Language (DDL)
- ۲- دستورات عملیات روی داده‌ها (Data Manipulation Language (DML)
- ۳- دستورات کنترل داده‌ها (Data Control Language (DCL)

تقسیم‌بندی زبان داده‌ای فرعی از نظر نیاز به زبان میزبان:

زبان داده ای فرعی از نظر نیاز یا عدم نیاز به زبان میزبان به دو رده تقسیم می شود:

۱- مستقل (I.DSL):

به زبان میزبان نیاز ندارد و به صورت تعاملی استفاده می‌شود. در واقع یک زبان پرس وجو است یعنی با دستوراتش می تواند پرسشهایی برای انجام عملیات بازیابی , درج , حذف و بهنگام سازی را تنظیم کند.

۲- ادغام‌شدنی (E.DSL):

دستورهایش در متن برنامه‌ای به زبان میزبان به کار می‌رود و مستقلاً قابل استفاده نیست.

نکته:

زبان داده ای فرعی ممکن است هم مستقل و هم ادغام‌شدنی باشد. (I/E.DSL)

برخی ویژگیهای زبان داده‌ای فرعی:

- ۱- تعداد دستورهایش باید کم باشد.
- ۲- دستورهایش باید شبیه زبان طبیعی باشد.
- ۳- یادگیری و استفاده آن باید ساده باشد.
- ۴- در طراحی آن باید اصل وحدت دستور رعایت شود.
- ۵- دستورهایش باید مبتنی بر عناصر ساختاری اساسی ساختار داده‌ای طراحی شوند.
- ۶- بهتر است نارویه‌ای (ناروشمند) باشند.
- ۷- می‌تواند کامپایلری یا مفسری باشد.
- ۸- بهتر است از نظر ساختاری کامل باشد.
- ۹- باید از نظر برنامه‌سازی و محاسباتی کامل باشد.
- ۱۰- زبان باید از نظر تعداد دستورهای کنترل داده‌ها و عملکرد هردستور، غنی و قوی باشد.
- ۱۱- باید از نظر انواع داده‌ای و به ویژه انواع داده‌ای انتزاعی و انواع داده‌ای پیچیده غنی باشد.
- ۱۲- باید مبتنی بر اصل دو اسلوبی (دو راه و روشی) باشد.

جلسه هشتم

سیستم مدیریت پایگاه داده ها

Database Management System

آنچه در این جلسه می خوانید:

۱- سیستم مدیریت پایگاه داده‌ها (DBMS)

۲- امکاناتی که DBMS به کاربر می‌دهد

۳- رده‌بندی سیستم‌های مدیریت پایگاه داده‌ها

۴- اجزای DBMS از نمای بیرونی

۵- نمای بیرونی (ساده شده) DBMS

۶- اجزای DBMS از نمای درونی

۷- واحدهای لایه هسته

۸- واحدهای لایه مدیریت محیط پایگاه داده‌ها

۹- ساختار یک سیستم پایگاه‌های

۱۰- جایگاه سیستم مدیریت پایگاه داده‌ها

سیستم مدیریت پایگاه داده‌ها (DBMS):

یکی از انواع نرم‌افزارهای واسط بین محیط فیزیکی ذخیره و بازیابی اطلاعات و محیط منطقی برنامه‌سازی است که به کاربر امکان می‌دهد تا:

- ۱- پایگاه داده‌های خود را تعریف کند.
- ۲- در پایگاه داده‌های خود عملیات انجام دهد.
- ۳- روی پایگاه داده‌های خود تا حدی کنترل داشته باشد.

رده‌بندی سیستم‌های مدیریت پایگاه داده‌ها:

- ۱- از نظر نوع مدل داده
- ۲- از نظر محیط سخت افزاری
- ۳- از نظر رده کامپیوتر
- ۴- از نظر محیط سیستم عامل
- ۵- از نظر معماری سیستم پایگاه داده ها
- ۶- از نظر معماری مشتری - خدمت‌گزار
- ۷- از نظر زبان
- ۸- از نظر زبان داده ای فرعی
- ۹- از نظر ماهیت زبان داده ای فرعی
- ۱۰- از نظر سیستم فایل
- ۱۱- از نظر نوع کاربرد
- ۱۲- از نظر قیمت
- ۱۳- از نظر طرز برپایی
- ۱۴- از نظر واسط کاربر
- ۱۵- از نظر رفتار در قبال رویدادها
- ۱۶- از نظر متدولوژی زبان پایگاهی
- ۱۷- از نظر بهینه سازی پرسش
- ۱۸- از نظر تراکنش
- ۱۹- از نظر نوع پردازش
- ۲۰- از نظر رسانه ذخیره سازی پایگاه داده ها
- ۲۱- از نظر قابلیت تعامل بین سیستمها
- ۲۲- از نظر پردازش داده های زمانمند
- ۲۳- از نظر نسل تکنولوژی

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر نوع ساختار داده‌ای:

- سیستم رابطه‌ای
- سیستم سلسله‌مراتبی
- سیستم شبکه‌ای
- بجز اینها

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر محیط سخت‌افزاری:

- وابسته به یک محیط خاص
- ناوابسته به یک محیط خاص

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر رده کامپیوتر:

- خاص محیط کامپیوترهای شخصی (Personal Computer)
- خاص محیط کامپیوترهای متوسط (Mini Computer)
- خاص محیط کامپیوترهای بزرگ (Main Computer)
- خاص محیط کامپیوترهای خیلی بزرگ (Super Computer)
- اجرا شونده در چند رده کامپیوتر

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر محیط سیستم عامل:

- وابسته به یک سیستم عامل خاص
- اجرا شونده در محیط چند سیستم عامل

رده‌بندی سیستم‌های مدیریت پایگاه داده‌ها:

از نظر نوع معماری سیستم پایگاه داده‌ها:

- با توان ایجاد پایگاه متمرکز
- با توان ایجاد پایگاه نامتمرکز

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر معماری مشتری - خدمت‌گزار:

- با توان ایجاد معماری چند مشتری - یک خدمت‌گزار
- با توان ایجاد معماری چند مشتری - چند خدمت‌گزار

رده‌بندی سیستم‌های مدیریت پایگاه داده‌ها:

از نظر زبان:

- سیستم دارای SQL
- سیستم فاقد SQL

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر نوع زبان داده‌ای فرعی:

- دارای I.DSL
- دارای E.DSL
- دارای E/I.DSL

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر ماهیت زبان داده‌ای فرعی:

- با زبان رویه‌ای
- با زبان نارویه‌ای

زبان رویه‌ای، زبانی است که در برنامه سازی با آن، برنامه ساز خودش رویه یا روش انجام کار مورد نظر را مشخص می کند.

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر سیستم فایل:

- خودکفا
- وابسته به سیستم فایل محیط سیستم عامل

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر نوع کاربرد:

- تک منظوره
- همه منظوره

رده‌بندی سیستم‌های مدیریت پایگاه داده‌ها:

از نظر قیمت:

- از حدود ده هزار دلار تا صد هزار دلار و گاه بیشتر

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر طرز برپایی:

- با محدودیت برپایی یکپارچه
- دارای امکان برپایی گزینشی

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر واسط کاربر:

- با واسط زبانی
- با واسط غیر زبانی
- با هر دو واسط

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر رفتار در قبال رویدادها:

- سیستم مدیریت فعال
- سیستم مدیریت غیر فعال

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر متدولوژی زبان:

- بدون متدولوژی شیء‌گرایی
- دارای متدولوژی شیء‌گرایی

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر بهینه سازی پرسش:

- دارای بهینه‌سازی متعارف
- دارای بهینه‌سازی مبتنی بر قاعده، معنایی و ...

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر نوع تراکنش:

- پذیرنده تراکنشهای ساده و تک سطحی
- پذیرنده تراکنشهای با مدل پیشرفته (مثلاً تودرتو، زنجیره ای و ...)

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر نوع پردازش:

- با قابلیت پردازش بی‌درنگ
- فاقد این قابلیت

رده‌بندی سیستم‌های مدیریت پایگاه داده‌ها:

از نظر رسانه ذخیره‌سازی پایگاه داده‌ها:

- با قابلیت ایجاد پایگاه داده‌های مقیم در حافظه
- فاقد این قابلیت

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر قابلیت تعامل بین سیستم‌ها:

- فاقد این قابلیت
- دارای قابلیت تعامل (با سیستم‌های همگن و ناهمگن یا فقط همگن)

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر پردازش داده‌های زمانمند:

- فاقد جنبه‌های یک سیستم مدیریت زمانی (سیستم معمولی)
- سیستم مدیریت پایگاه داده‌های زمانی

رده‌بندی سیستم‌های مدیریت پایگاه داده ها:

از نظر نسل تکنولوژیک:

- نسل سیستم‌های پیش رابطه ای
- نسل سیستم‌های رابطه ای
- نسل سیستم‌های پسا رابطه ای (شیء گرا و شیء رابطه ای)

اجزای سیستم مدیریت پایگاه داده ها:

سیستم مدیریت پایگاه داده ها شامل دو دید (نما) است:

۱- نمای بیرونی

۲- نمای درونی

سیستم مدیریت پایگاه داده ها (نمای بیرونی):

کاربر

برنامه‌های کاربردی / پرسشها

سیستم مدیریت پایگاه داده‌ها

واحد پردازشگر برنامه کاربردی و پرسشها

واحد دستیابی به داده‌های ذخیره‌شده

DB

داده‌های
ذخیره‌شده

مَتا داده

سیستم مدیریت پایگاه داده ها از نمای
بیرونی از دو واحد اصلی تشکیل شده است:

- ۱- واحد پردازشگر پرسش‌ها و برنامه‌های کاربردی
- ۲- واحد ایجاد و مدیریت داده‌های ذخیره‌شده

سیستم مدیریت پایگاه داده ها (نمای درونی):

سیستم مدیریت پایگاه داده ها از نمای درونی از سه لایه تشکیل شده است:

۱- لایه هسته (سیستم کنترل یا موتور پایگاه داده ها)

۲- لایه مدیریت محیط پایگاه داده ها

۳- لایه تسهیلات نرم افزاری (ابزارها)

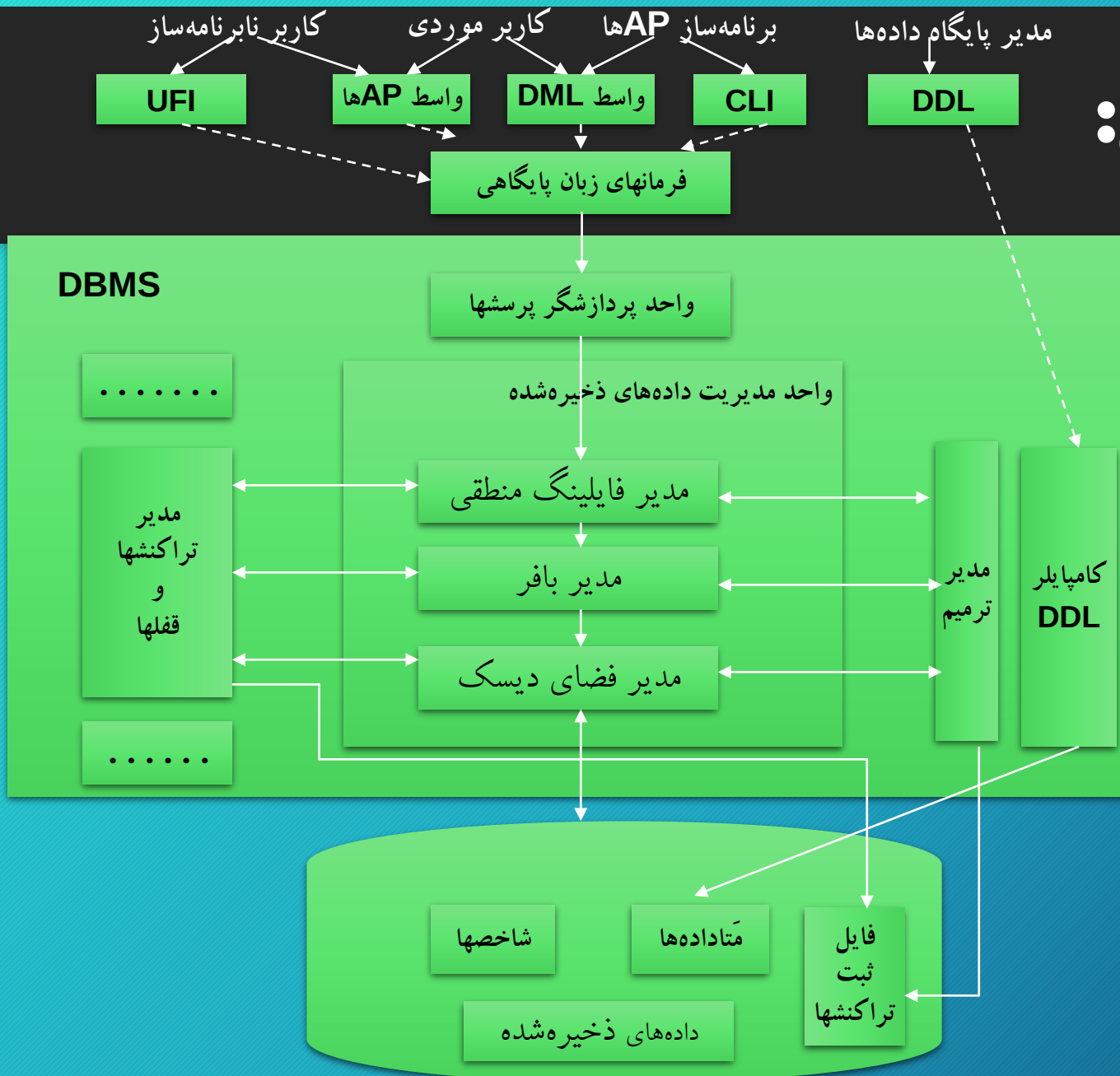
واحدهای لایه هسته:

- ۱- واحد دریافت درخواست کاربر و واریسی‌های اولیه
- ۲- واحد تولید شِماها
- ۳- پیش‌کامپایلرها برای DML
- ۴- کامپایلرها (پردازنده DML)
- ۵- پردازشگر پرسش و بهینه‌ساز پرسش
- ۶- واحد مدیریت سطح داخلی
- ۷- واحد مدیریت بافر
- ۸- واحد مدیریت فضای دیسک
- ۹- واحد ناظر زمان اجرا
- ۱۰- واحد مدیریت همروندی تراکنش‌ها
- ۱۱- واحد مدیریت انتقال داده‌ها
- ۱۲- واحد مدیریت کاتالوگ

واحدهای لایه مدیریت محیط پایگاه داده ها:

- ۱- واحد کنترل جامعیت پایگاه داده ها
- ۲- واحد ترمیم (بازسازی) پایگاه داده ها
- ۳- واحد ایمنی و حفاظت پایگاه داده ها
- ۴- واحد تولید نسخه های پشتیبان
- ۵- واحد تولید فایل های ثبت تراکنشها

ساختار سیستم پایگاهی:



جایگاه سیستم مدیریت پایگاه داده ها در یک سیستم کامپیوتری:



یک DBMS برای هر برنامه کاربردی:

درخواست کاربر

سیستم عامل

AP1

AP_i

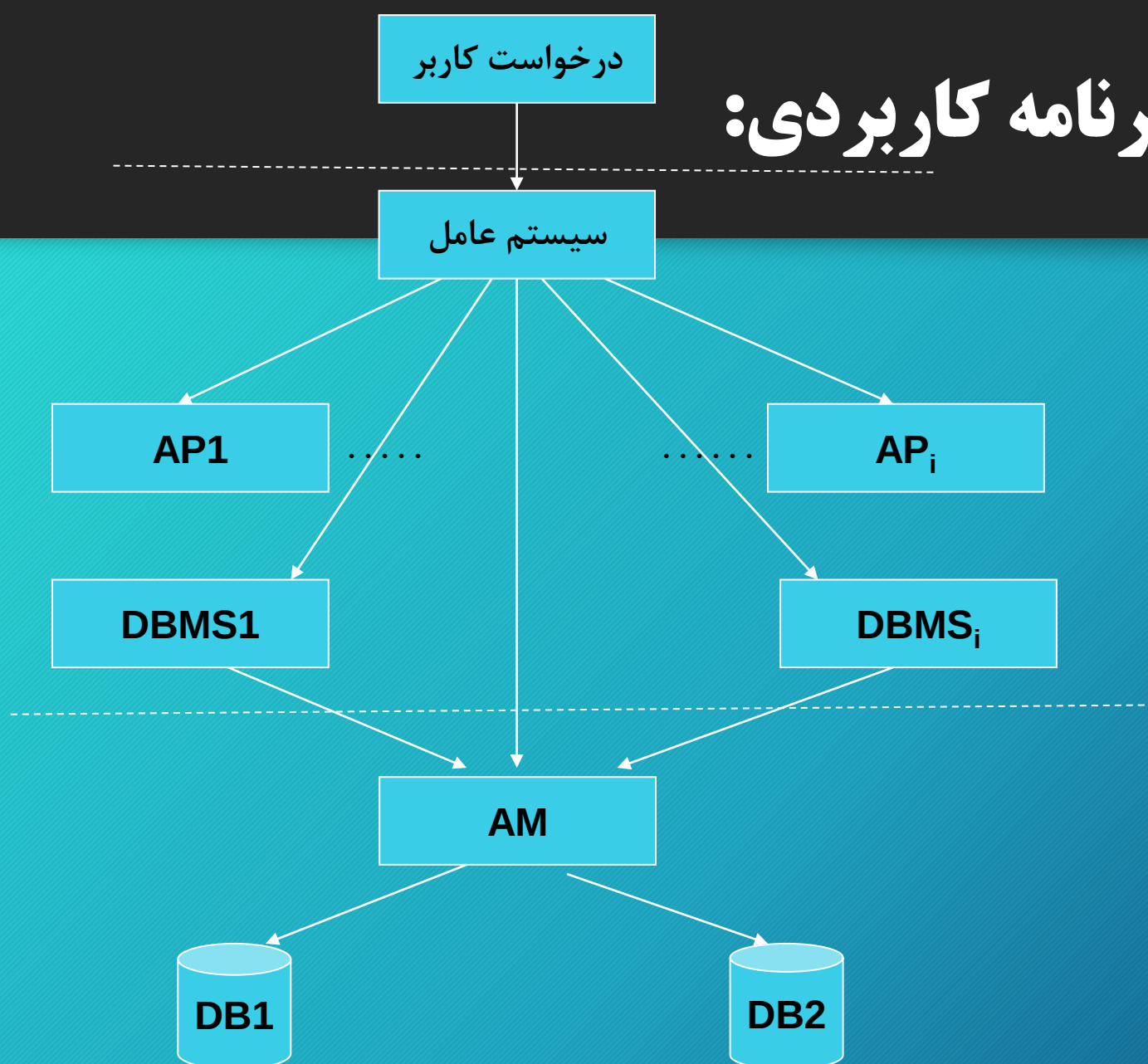
DBMS1

DBMS_i

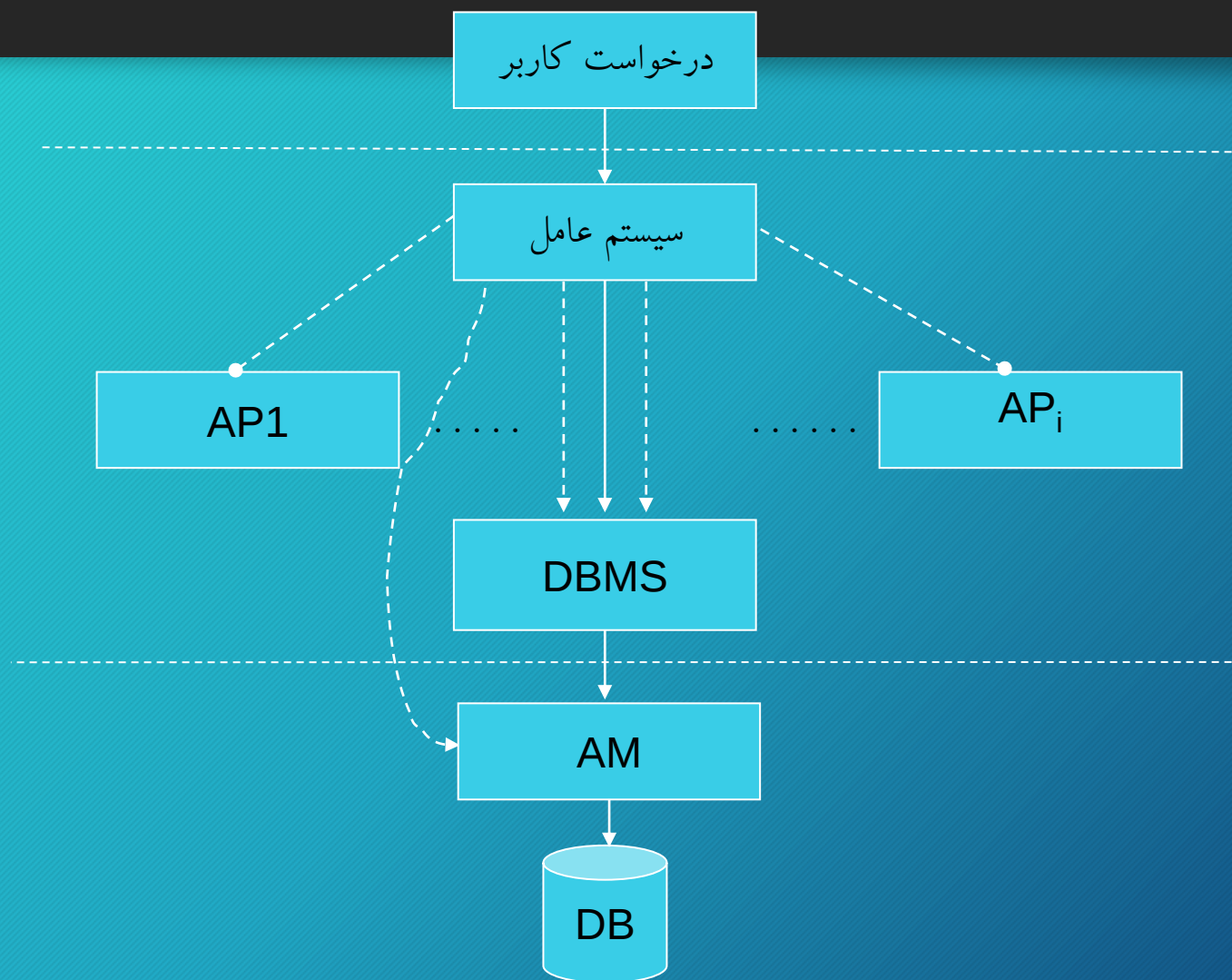
AM

DB1

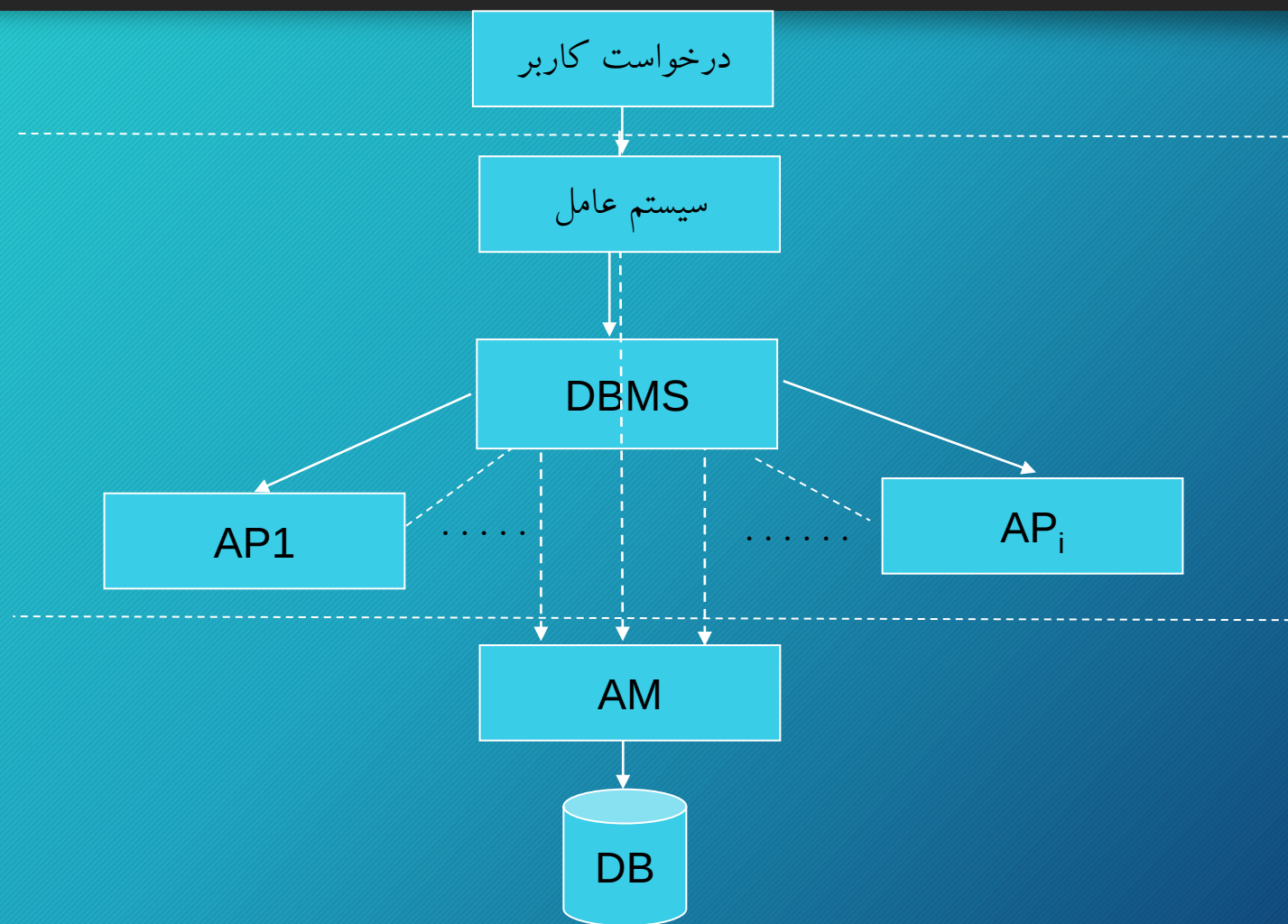
DB2



یک DBMS برای چند برنامه کاربردی تحت کنترل سیستم عامل:



یک سیستم مدیریت برای چند برنامه کاربردی با اجرای تحت کنترل DBMS:



روشها, حالتها یا اسلوبهای عملیاتی:

• به طور کلی سه اسلوب عملیاتی وجود دارد:

۱- اسلوب یکجا

۲- اسلوب برخط

۳- اسلوب تعاملی

کاتالوگ سیستم و دیکشنری داده‌ها (مَتا داده‌ها):

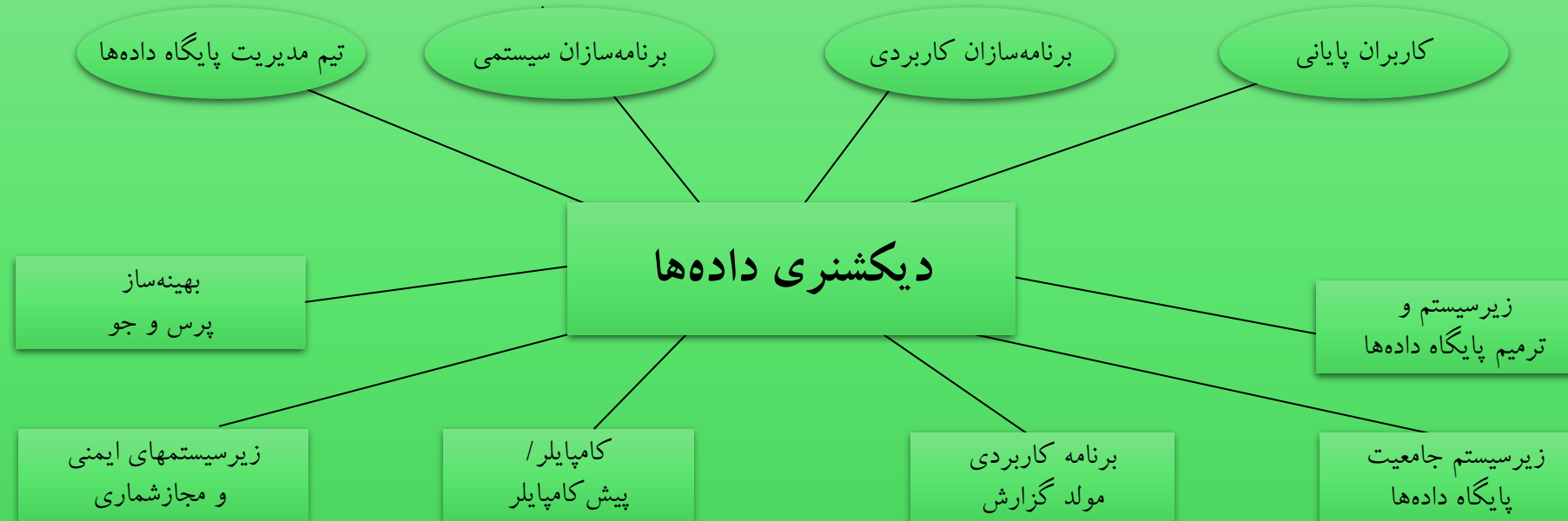
حاوی داده‌هایی است در مورد داده‌های ذخیره‌شده در پایگاه داده‌های کاربر و این داده‌ها به مَتاداده‌ها موسومند.

دیکشنری داده‌ها معمولاً جزئی از خود سیستم است و به دو صورت فعال و غیرفعال تولید می‌شود.

اطلاعاتی که در دیکشنری داده‌ها نگهداری می‌شود:

- شمای خارجی
- شمای ادراکی
- شمای داخلی
- رویه‌های مربوط به تبدیلات بین سه سطح معماری
- شرح ساختار فیزیکی داده‌های ذخیره‌شده
- مشخصات و حقوق دستیابی کاربران به داده‌ها
- مشخصات برنامه‌های کاربردی تولید شده و ارتباط آنها با درخواستهای کاربران
- مشخصات پایانه‌های متصل به سیستم
- ارتباط بین برنامه‌های کاربردی و داده‌های ذخیره‌شده
- قواعد جامعیت
- ضوابط کنترل ایمنی داده‌ها
- مشخصات پیکربندی سخت‌افزاری سیستم و رسانه‌های ذخیره‌سازی
- اطلاعات متنوع آماری پایگاه داده و کاربران
- توابع تعریف‌شده توسط کاربران

دیکشنری داده‌ها و استفاده‌کنندگان آن:



موارد مهم دیگر در سیستم مدیریت پایگاه داده ها:

- ۱- مشخصات کلی سیستم
- ۲- تسهیلات و جنبه های دیگر سیستم
- ۳- پارامترهای مربوط به معماری پایگاه داده ها
- ۴- پارامترهای مربوط به زبان داده ای فعلی
- ۵- محورهای اصلی مقایسه سیستمهای مدیریت
- ۶- طرز مطالعه سیستم مدیریت
- ۷- رویه های مستند برای کاربران
- ۸- هزینه ها
- ۹- روش ارزیابی

...

مدیر پایگاه داده ها

مزایا و معایب سیستم پایگاه‌های

آنچه در این جلسه می خوانید:

- ۱- مدیر پایگاه داده ها
- ۲- برخی از مسئولیتهای تیم مدیریت پایگاه داده
- ۳- سرمایه های سازمان در مدیریت نوین
- ۴- مزایا و معایب سیستم پایگاهی تک کاربری
- ۵- مزایا و معایب سیستم پایگاهی چند کاربری
- ۶- شرایط استفاده از تکنولوژی پایگاه داده ها
- ۷- کاربردهای جدید تکنولوژی پایگاه داده ها

مدیر پایگاه داده‌ها:

فردی است متخصص در پایگاه داده‌ها و با مسئولیت علمی، فنی و نیز اداری در محدوده وظایفی که عهده‌دار است.

این مدیر همراه با یک تیم تخصصی کار می‌کند که به آن تیم مدیریت پایگاه داده‌ها می‌گویند. هر یک از اعضای این تیم مسئولیت خاصی دارد و در حیطه اختیارات و وظایفش می‌تواند سرپرست یک تیم اجرایی باشد.

برخی از مسئولیتهای اصلی در تیم مدیریت پایگاه داده‌ها:

- مدیر پایگاه داده ها

- مسئول کنترل کارایی DBMS

- مدیر داده ها

- مسئول کنترل کارایی سیستم پایگاهی

- مدیر امور پژوهش - توسعه

- مسئول نظارت بر عملیات روی پایگاه داده‌ها

- مدیر سیستمهای کاربردی

- مسئول تماس با کاربران زیرمحیطهای سازمان

- مسئول تیمهای برنامه سازی

- مسئول تنظیم مستندات و وضع استانداردها

- مسئول ایمنی و حفاظت سیستم پایگاهی

در مدیریت نوین سازمانها، هر سازمان دارای پنج سرمایه است:

▪ سخت افزار

▪ نرم افزار

▪ داده

▪ بودجه

▪ تخصص

مزایا و معایب سیستم پایگاهی:

مزایای این تکنولوژی بستگی به عواملی از جمله: نوع سیستم مدیریت پایگاه داده ها, معماری سیستم پایگاهی و ماهیت کاربردها دارد.

مزایا و معایب سیستم پایگاهی را از نظر نوع کاربری می توان بررسی کرد.

مزایا و معایب سیستم پایگاهی تک کاربری:

مزایا:

- هر بخش از سازمان، داده‌های خود را نگهداری و پردازش می‌کند.
- با استفاده از کامپیوترهای شخصی، حجم داده‌های سیستم مرکزی کاهش می‌آید.
- پایگاه داده‌های ایجادشده روی کامپیوترهای شخصی معمولاً کوچک و مدلسازی، طراحی و پیاده‌سازی آنها ساده است.
- کار با این سیستمها و برنامه‌سازی در محیط آنها ساده است.
- با پیشرفت کامپیوترهای شخصی، این سیستمها می‌توانند بسیاری از کارهای سیستمهای کامپیوتری بزرگ را انجام دهند.

مزایا و معایب سیستم پایگاهی تک کاربری:

معایب:

- وجود تعداد زیادی سیستم پایگاهی کوچک در یک سازمان باعث بروز افزونگی، ناسازگاری داده‌ها و نایمنی آنها می‌شود.
- محدودیتهای سخت‌افزاری سبب محدودیت اندازه فایلها و نیز محدودیت سرعت پردازش باعث محدود شدن حجم پایگاه داده‌ها می‌شود.
- خود سیستم نمی‌تواند قوی و کارا باشد.
- میزان ایمنی و حفاظت در آنها ضعیف است.
- امکانات تولید نسخه پشتیبان در آنها معمولاً کم است.
- اشتراکی کردن آنها مشکلات تکنیکی جدی دارد.
- اِعمال مجموعه واحدی از استانداردها در کل سازمان ناممکن است.
- معمولاً کاربر این محیط مهارت کافی در مدلسازی و طراحی بهینه پایگاه داده‌ها ندارد.
- ایجاد یک سیستم جامع و یکپارچه، براساس این سیستمها، دشوار و پرهزینه است.

مزایا و معایب سیستم پایگاهی چند کاربری:

مزایا:

- اشتراک داده‌ها
- کاهش افزونگی
- تعدّد و تسهیل شیوه‌های دستیابی به داده‌ها
- اجتناب از ناسازگاری داده‌ها
- تأمین همروندی بهتر
- تسهیل پردازش تراکنش‌ها
- تضمین جامعیت داده‌ها
- حفظ محرمانگی داده‌ها
- امکان اعمال استاندارد‌ها
- تعدد زبانها
- کاهش حجم برنامه‌ها
- تنوع کاربران
- تسريع در دریافت پاسخ پرسش‌ها
- استفاده بهتر از سخت‌افزار و ...

مزایا و معایب سیستم پایگاهی چند کاربری:

معایب:

- هزینه بالای نرم افزار و سخت افزار
- هزینه بیشتر برای برنامه سازی
- هزینه بالا برای انجام مهندسی دوباره
- گند شدن اجرای بعضی از برنامه های کاربردی
- خطر آسیب پذیری داده ها
- تأثیرات گسترده تر خرابیها و دشواری بیشتر ترمیم آنها
- پیچیده بودن سیستم و نیاز به تخصص بیشتر

شرایط استفاده از تکنولوژی پایگاه داده‌ها:

از دیدگاه تکنیکی و کاربردی چنانچه شرایط زیر برقرار باشد، بکارگیری تکنولوژی پایگاه داده‌ها در سازمان توصیه می‌شود:

- نیاز به ایجاد یک سیستم یکپارچه اطلاعاتی
- حجم زیاد داده‌های سازمان و رشد پویای آن
- تغییرات مداوم در داده‌های ذخیره‌شده
- بالا بودن بسامد درخواستهای کاربران
- نیاز به اعمال کنترل متمرکز و دقیق روی کل داده‌ها
- وجود ارتباطات پیچیده بین داده‌ها
- زیاد بودن میزان داده‌های مشترک بین برنامه‌های کاربردی
- مدنظر بودن صحت، دقت و سازگاری داده‌ها
- زیاد بودن گزارشها
- نیاز به انجام پردازشهای تحلیلی برخط
- نیاز به سیستم داده‌کاوی و کشف دانش در سازمان

کاربردهای جدید تکنولوژی پایگاه داده‌ها:

- سیستم پشتیبان تصمیم
- سیستم داده‌کاوی چندرسانه‌ای
- سیستم انبارش داده‌ها
- حیطه‌های فضایی و جغرافیایی
- پایگاه داده‌های شخصی و همراه
- سیستم پایگاه داده‌ها در شبکه جهانی
- اطلاع‌رسانی
- سیستم اطلاعات اجرایی
- سیستم اطلاعات طراحی
- سیستم پردازش تحلیلی برخط چندبعدی
- ...

جلسه دهم

مفاهیم اساسی مدل داده رابطهای

آنچه در این جلسه می خوانید:

- ۱- تعریف رابطه
- ۲- چکیده و درجه و کاردینالیتی رابطه
- ۳- تناظر بین مفاهیم رابطه‌ای و اصطلاحات جدولی
- ۴- ویژگیهای رابطه
- ۵- انواع رابطه
- ۶- میدان (دامنه)
- ۷- مزایای میدان
- ۸- رابطه نرمال و غیرنرمال
- ۹- مزایا و معایب رابطه نرمال و غیرنرمال
- ۱۰- انواع کلید در مدل رابطه‌ای
- ۱۱- قواعد جامعیت در مدل رابطه‌ای و انواع آن
- ۱۲- بخشهای مدل داده در مدل داده رابطه‌ای
- ۱۳- مزایا و معایب مدل رابطه‌ای
- ۱۴- کاتالوگ در مدل رابطه‌ای

تعریف رابطه:

(A) نام صفت (V) مقدار صفت

نام خانوادگی احمدی
ش. دانشجویی ۹۶۱۰۷۷۵۲۱

هر رابطه از دو مجموعه تشکیل شده است:

۱- عنوان (سر آیند):

مجموعه‌ای است نامدار از n عنصر به صورت $\langle D_i : A_i \rangle$ ($i=1, 2, 3, \dots, n$) که A_i نام یک صفت و D_i نام میدان صفت است.

$\{\langle D_1 : A_1 \rangle, \langle D_2 : A_2 \rangle, \dots, \langle D_n : A_n \rangle\}$

۲- بدنه (پیکر):

مجموعه‌ای است از تعدادی عنصر, هر یک مجموعه‌ای از سه تایی‌های مرتب به صورت $\langle D_i : A_i : v_{ij} \rangle$ ($i, j=1, 2, 3, \dots, n$) که v_{ij} یک مقدار برای صفت A_i است.

چکیده و درجه و کاردینالیتی رابطه:

چکیده رابطه:

به نام و مجموعه صفات رابطه, چکیده رابطه می گویند.

$$R(A_1, A_2, \dots, A_n)$$

درجه رابطه (آریتی رابطه):

به تعداد صفات یک رابطه درجه رابطه می گویند.

کاردینالیتی رابطه:

تعداد سطرهای (تاپلها) رابطه را کاردینالیتی می گویند.

تناظر بین مفاهیم رابطه‌ای و اصطلاحات جدولی:

اصطلاح جدولی	مفهوم رابطه‌ای
جدول	رابطه
سطر	تاپل
ستون	صفت
مجموعه مقادیر ستون	میدان
تعداد ستونها	درجه (آریتی)
تعداد سطرها	کاردینالیتی

نام	نام خانوادگی	شماره دانشجویی	شماره ملی
علی	احمدی	۹۵۱۲۰۴۴۸۶	۰۹۹۴۲۵۳۲۱۴
محمد	حیدری	۹۷۵۶۳۲۱۴۷	۰۶۶۱۵۸۴۳۶۵
ریحانه	محمدزاده	۹۸۲۵۳۶۴۸۸	۰۸۵۵۶۸۶۲۲۲

ویژگیهای رابطه:

- ۱- رابطه تاپل تکراری ندارد.
- ۲- تاپلها نظم ندارند.
- ۳- صفات رابطه نظم مکانی ندارند.
- ۴- تمام صفات رابطه تک مقداری هستند.

انواع رابطه:

▪ رابطه نامدار:

رابطه‌ای است با یک نام مشخص.

▪ رابطه مبنا:

رابطه‌ای نامدار که استقلال وجودی دارد.

▪ رابطه مجازی:

نوعی رابطه نامدار که مشتق از رابطه‌های دیگر است.

▪ رابطه لحظه‌ای:

نوعی رابطه نامدار و مشتق که مجازی نیست.

▪ رابطه مشتق:

رابطه‌ای که به کمک یک عبارت رابطه‌ای تعریف می‌شود.

▪ رابطه عبارتی:

نوعی رابطه که با یک عبارت رابطه‌ای بدست می‌آید.

▪ رابطه نتیجه پرسش:

رابطه‌ای بی‌نام و مشتق، حاصل اجرای یک پرسش مشخص.

▪ رابطه بینابینی:

رابطه‌ای بی‌نام و مشتق، حاصل ارزیابی یک عبارت درون یک عبارت بزرگتر.

▪ رابطه موقت:

رابطه‌ای نامدار که استقلال وجودی دارد ولی داده هایش پایدار نیستند.

میدان (دامنه):

میدان مجموعه‌ای است نامدار و متناهی از مقادیر هم‌نوع و شناخته شده.

نکته:

چون رابطه روی تعدادی میدان تعریف می‌شود، برای تعریف و ایجاد یک رابطه منطقاً باید ابتدا میدانهایش را تعریف کنیم.

مزایای میدان:

- ۱- سبب ساده‌تر شدن و کوتاه‌تر شدن شمای پایگاه داده‌ها می‌شود.
- ۲- تغییر در شمای پایگاه را تسهیل می‌کند.
- ۳- امکانی است برای کنترل مقداری عملیات در پایگاه داده‌ها.
- ۴- امکانی است برای کنترل معنایی درخواستها (پرسشها).
- ۵- پاسخگویی به برخی از پرسشها را آسان می‌کند.
- ۶- دگرنامی صفات را تسهیل می‌کند.
- ۷- اعمال قواعد جامعیت را تسهیل می‌کند.
- ۸- با استفاده از مفهوم میدان می‌توان امکانات و جنبه‌های موجود در سیستم مدیریت پایگاه داده‌های شیء‌گرا را به سیستم مدیریت پایگاه داده‌های رابطه‌ای افزود.

رابطه نرمال و غیر نرمال:

رابطه نرمال:

رابطه‌ای است که تمام صفات آن تک‌مقداری باشند.

رابطه غیر نرمال:

رابطه‌ای است که حداقل یک صفت (ساده یا مرکب) چندمقداری داشته باشد.

مزایا و معایب رابطه نرمال:

مزایای رابطه نرمال:

- ۱- سادگی در نمایش ظاهری رابطه.
- ۲- سادگی دستورات.
- ۳- سادگی در اجرای عملیات در پایگاه داده‌ها.

معایب رابطه نرمال:

- ۱- بروز پدیده افزونگی.
- ۲- طولانی تر شدن کلید رابطه.
- ۳- دشواری در نمایش داده‌های پیچیده.
- ۴- دشواری در نمایش طبیعی مفهوم سلسله مراتب.
- ۵- دشواری در نمایش مفهوم ارث‌بری.

مزایا و معایب رابطه غیر نرمال:

مزایای رابطه غیر نرمال:

- ۱- کاهش میزان افزونگی.
- ۲- کوتاه شدن کلید.
- ۳- امکان نمایش داده‌های پیچیده.
- ۴- دشواری کمتر در مفهوم سلسله مراتب و ارث‌بری.
- ۵- افزایش سرعت بازیابی در بعضی از پرسش‌ها.
- ۶- عدم نیاز به نرمالترسازی رابطه.

معایب رابطه غیر نرمال:

- ۱- پیچیدگی.
- ۲- عدم تقارن صفات.

انواع کلید در مدل رابطه‌ای:

۱- سوپر کلید (آبر کلید):

هر زیر مجموعه از عنوان رابطه که یکتایی مقدار داشته باشد.

۲- کلید کاندید:

هر زیر مجموعه از عنوان رابطه که خاصیت یکتایی مقدار و کاهش ناپذیری داشته باشد.

۳- کلید اصلی:

یکی از کایدهای کاندید که از نظر کاربر رایج بوده و مقادیرش معلوم بوده و تغییر نکند و طولش کوتاه باشد.

۴- کلید بدیل (دیگر):

هر کلید کاندید غیر از کلید اصلی.

۵- کلید خارجی:

در دو رابطه هر زیرمجموعه از صفات که مقادیرشان برابر باشند. (برای نمایش ارتباط بین موجودیتها)

قواعد جامعیت در مدل رابطه‌ای:

جامعیت پایگاه داده‌ها یعنی صحت، سازگاری، دقت و اعتبار داده‌های ذخیره شده در پایگاه در تمام لحظات.

عوامل نقض جامعیت:

- اشتباه در وارد کردن داده‌ها.
- اشتباه در برنامه‌های کاربردی.
- وجود افزونگی کنترل نشده.
- وارد کردن تراکنشها به گونه‌ای که داده نامعتبر ایجاد شود.
- خرابیهای سخت‌افزاری و نرم‌افزاری.

انواع قواعد جامعیت در مدل رابطه‌ای:

■ قواعد کاربری (قواعد خاص):

این قواعد, قواعدی هستند که توسط کاربر مجاز تعریف می‌شوند.

■ مَتا قواعد (قواعد عام):

این قواعد, قواعدی هستند که باید توسط هر سیستم رابطه‌ای در هر پایگاه داده‌های رابطه‌ای اعمال شوند.

انواع قواعد کاربری (قواعد خاص):

■ محدودیت میدانی:

نوع و مقادیر آنرا مشخص می کند.

■ محدودیت صفتی (ستونی):

نوع و مقادیر آنرا مشخص می کند.

■ محدودیت رابطه‌ای:

مقادیر مجاز یک رابطه را مشخص می کند.

■ محدودیت پایگاهی:

ناظر است به دو یا بیش از دو متغیر رابطه‌ای که به نحوی با هم ارتباط معنایی دارند.

انواع متاقواعد:

■ قاعده C_1 (قاعده جامعیت موجودیتی):

هیچ جزء تشکیل دهنده کلید اصلی نمی تواند هیچ مقدار پذیر باشد.

■ قاعده C_2 (قاعده جامعیت ارجاعی):

اگر یک صفت در یک رابطه کلید خارجی باشد، آن صفت در آن رابطه می تواند مقدار نداشته باشد به شرطی که جزء تشکیل دهنده کلید آن رابطه نباشد، در غیر اینصورت حتماً باید مقداری داشته باشد که در رابطه مرجع وجود داشته باشد.

بخشهای مدل داده در مدل داده رابطه‌ای:

مدل رابطه‌ای یک مدل داده کامل است زیرا هر سه بخش اساسی مدل داده به شرح زیر را دارد:

▪ بخش ساختاری:

رابطه و عناصر ساختاری‌اش یعنی: تاپل، صفت، میدان، درجه، کاردینالیتی، هیچمقدار و کلیدها.

▪ بخش برزشی (پردازشی):

امکانات جبر رابطه‌ای و حساب رابطه‌ای.

▪ بخش جامعیتی:

▪ قواعد جامعیت کاربری، قاعده C_1 و قاعده C_2 .

مزایای مدل رابطه‌ای:

- ۱- نمایش ساده.
- ۲- مبنای ریاضی دارد.
- ۳- تامین‌کننده استقلال داده‌ای است.
- ۴- یک عنصر ساختاری اساسی به نام رابطه دارد.
- ۵- کامل است.
- ۶- عدم نیاز به مفهوم کلاسیک نظم.
- ۷- برای بهترسازی طراحی، ابزار طراحی بهتر دارد.
- ۸- ...

معایب مدل رابطه‌ای:

- ۱- دشواری در نمایش داده‌های پیچیده.
- ۲- دشواری در نمایش مفاهیمی مانند وراثت، تعمیم و ...
- ۳- نداشتن قواعد جامعیت ذاتی.
- ۴- عدم امکان مدلسازی رفتار اشیاء.
- ۵- فقدان اکمال برنامه‌سازی و اکمال محاسباتی.
- ۶- دشواری در یافتن کلید کاندید.
- ۷- جدا بودن داده‌ها از روشها.
- ۸- مشکلات ناشی از وجود مفهوم هیچ مقدار.
- ۹- ...

کاتالوگ در مدل رابطه‌ای:

کاتالوگ حاوی داده‌هایی است در مورد داده‌های ذخیره شده در پایگاه داده‌ها. معمولاً در سیستم‌های رابطه‌ای دو نوع کاتالوگ ایجاد می‌شود:

کاتالوگ کاربر

کاتالوگ DBA

اطلاعاتی که در کاتالوگ نگهداری می‌شوند به شرح زیر می‌باشند:

- میدان
- رابطه
- صفت
- دید
- قواعد جامعیت
- کاتالوگ حاوی داده‌هایی است در مورد داده‌های ذخیره شده در پایگاه داده‌ها.
- توابع و رویه‌های تعریف‌شده توسط کاربر
- رهانا‌های تعریف‌شده توسط کاربر
- کاربر
- شاخص و ساختارهای مرتبط با آن
- ضوابط ایمنی
- پرسش‌ها
- برنامه‌های کاربردی

جبر رابطه‌ای:

عملیات در پایگاه دادهٔ رابطه‌ای

آنچه در این جلسه می خوانید:

- ۱- عملیات در پایگاه دادهٔ رابطه‌ای
- ۲- جبر رابطه‌ای
- ۳- عملگرهای معمولی و خاص
- ۴- عملگر اجتماع
- ۵- عملگر اشتراک
- ۶- عملگر تفاضل
- ۷- عملگر ضرب
- ۸- عملگر گزینش (تحدید)
- ۹- عملگر پرتو معمولی
- ۱۰- عملگر پرتو گسترش یافته
- ۱۱- عملگر پیوند
- ۱۲- عملگر تقسیم
- ۱۳- مثالهایی از کاربرد عملگرها
- ۱۴- کاربردهای جبر رابطه‌ای
- ۱۵- حساب رابطه‌ای

عملیات در پایگاه دادهٔ رابطه‌ای:

برای انجام عملیات در پایگاه دادهٔ رابطه‌ای به عملگرهای ریاضی نیاز داریم. برای این منظور دو امکان وجود دارد:

۱- جبر رابطه‌ای

۲- حساب رابطه‌ای

جبر رابطه‌ای:

برای عملیات در پایگاه داده رابطه‌ای از یکسری عملگرها استفاده می‌شود که این عملگرها به دو دسته زیر تقسیم می‌شوند:

۱- عملگرهای معمولی:

اجتماع اشتراک تفاضل ضرب کارتزین

۲- عملگرهای خاص:

گزینش (تحدید) پرتو پیوند تقسیم

عملگرهای معمولی

عملگر اجتماع:

R1 union R2

R1

نام خانوادگی	نام
احمدی	علی
حیدری	محمد

R2

نام خانوادگی	نام
غفاری	ریحانه

R1 U R2

نام خانوادگی	نام
احمدی	علی
حیدری	محمد
غفاری	ریحانه

عملگر اشتراک:

R1 intersect R2

R1

نام خانوادگی	نام
احمدی	علی
حیدری	محمد

R2

نام خانوادگی	نام
غفاری	ریحانه
حیدری	محمد

$R1 \cap R2$

نام خانوادگی	نام
حیدری	محمد

عملگر تفاضل:

R1 minus R2

R1

نام خانوادگی	نام
احمدی	علی
حیدری	محمد

R2

نام خانوادگی	نام
حیدری	محمد

R1 - R2

نام خانوادگی	نام
احمدی	علی

عملگر ضرب:

R1 product R2
R1 times R2

R1

نام خانوادگی	نام
احمدی	علی

R2

نام خانوادگی	نام
حیدری	محمد

R1 × R2

(علی, محمد)
(احمدی, حیدری)

عملگرهای خاص

عملگر گزینش (تحدید):

این عملگر تک عملوندی است و به صورتهای زیر نوشته می شود:

- R where شرط
- restrict R where شرط
- $\sigma_p(R)$
- $\sigma_{\text{شرط}}(R)$

عملکرد:

تاپلهایی (سطرهای) از رابطه R را برمی گرداند.

مثال عملگر گزینش (تحدید):

شرط where R

student where id='76124444'

نتیجه

نام خانوادگی	نام	id
احمدی	علی	76124444

student

نام خانوادگی	نام	id
احمدی	علی	76124444
حیدری	محمد	76258698

student where نام='محمد'

نتیجه

نام خانوادگی	نام	id
حیدری	محمد	76258698

مثال عملگر گزینش (تحدید):

σ (R) شرط

$\sigma_{\langle \text{id}='76124444' \rangle}$ (student)

نتیجه

id	نام	نام خانوادگی
76124444	علی	احمدی

student

id	نام	نام خانوادگی
76124444	علی	احمدی
76258698	محمد	حیدری

$\sigma_{\langle \text{نام}='محمد' \rangle}$ (student)

نتیجه

id	نام	نام خانوادگی
76258698	محمد	حیدری

عملگر پرتو:

▪ پرتو معمولی

▪ پرتو گسترش یافته

عملگر پرتو (پرتو معمولی):

این عملگر تک عملوندی است و به صورت‌های زیر نوشته می‌شود:

▪ $\text{project } R \text{ over } (A_1, A_2, \dots, A_i)$

▪ $\Pi_{\langle A_1, A_2, \dots, A_i \rangle} (R)$

عملکرد:

مقادیر صفاتی از رابطه R (از همه تاپل‌ها، سطرها) را برمی‌گرداند.

مثال عملگر پرتو (پرتو معمولی):

project R over($A_1, A_2, \dots, A_i$)

Project student over(id)



id
76124444
76258698

student

id	نام	نام خانوادگی
76124444	علی	احمدی
76258698	محمد	حیدری

Project student over(نام خانوادگی, نام)



نام خانوادگی	نام
احمدی	علی
حیدری	محمد

مثال عملگر پرتو (پرتو معمولی):

$\Pi_{\langle A_1, A_2, \dots, A_i \rangle} (R)$

$\Pi_{\langle id \rangle} (student)$

نتیجہ

id
76124444
76258698

student

id	نام	نام خانوادگی
76124444	علی	احمدی
76258698	محمد	حیدری

$\Pi_{\langle \text{نام خانوادگی, نام} \rangle} (student)$

نتیجہ

نام خانوادگی	نام
احمدی	علی
حیدری	محمد

عملگر پرتو (پرتو گسترش یافته):

این عملگر امکان می‌دهد تا در لیست صفات پرتو، از توابع حسابی استفاده شود.

شکل کلی این عملگر به صورت زیر است:

$$\Pi_{\langle F_1, F_2, \dots, F_i \rangle} (R)$$

عملکرد:

مانند پرتو معمولی

مثال عملگر پرتو (پرتو گسترش یافته):

$\Pi_{\langle F_1, F_2, \dots, F_i \rangle} (R)$

$\Pi_{\langle \text{نمره} + 0.1 \times \text{نمره} \rangle} (\text{student})$



نمره
۱۹.۸
۱۸.۷

student

id	نام	نمره
76124444	علی	۱۸
76258698	محمد	۱۷

عملگر پیوند:

این عملگر دو عملوندی است و به صورتهای زیر نوشته می شود:

- $R1 \bowtie R2$
- $R1 \text{ join } R2$
- $R1 \theta\text{-join } R2$
- $R1 \bowtie_{\text{شرط}} R2$

عملکرد:

رابطه جواب رابطه ای است که هر تاپل رابطه جواب از پیوند هر دو تاپل حائز شرط پیوند از دو رابطه $R1$ و $R2$ بدست می آید.

مثال عملگر پیوند:

R1 ⋈ R2

student ⋈_{c=گدرس s=گدرس} course

نتیجه

student

id	نام	نام خانوادگی	سگدرس
76124444	علی	احمدی	D222
76258698	محمد	حیدری	D444

course

نوع درس	تعداد واحد	نام درس	Cگدرس
تئوری	۳	ریاضی	D222
عملی	۳	کامپیوتر	D333

Id	نام	نام خانوادگی	گدرس	نام درس	تعداد واحد	نوع درس
76124444	علی	احمدی	D222	ریاضی	۳	تئوری

عملگر تقسیم:

این عملگر دو عملوندی است و به صورتهای زیر نوشته می شود:

R1 divide by R2

عملکرد:

حاصل شامل تاپلهایی مثل t به شرطی که تاپل t_{R1} , به ازای تمام تاپلهای t_{R2} در $R1$ وجود داشته باشد.

مثالهایی از کاربرد عملگرها:

- $\Pi_{\langle id \rangle}(\text{student}) \text{ union } \Pi_{\langle \text{نام درس} \rangle}(\text{course})$
- $\Pi_{\langle id \rangle}(\text{student}) \text{ intersect } \Pi_{\langle \text{نام درس} \rangle}(\text{course})$
- $\Pi_{\langle id \rangle}(\text{student}) \text{ minus } \Pi_{\langle \text{نام درس} \rangle}(\text{course})$
- ...

کاربردهای جبر رابطه‌ای:

- ۱- بازیابی داده‌ها
- ۲- ذخیره‌سازی داده‌ها
- ۳- تعریف انواع رابطه‌های مشتق
- ۴- تعریف قواعد برای کنترل پایگاه داده‌ها
- ۵- تعریف داده‌ها به عنوان حیطة بعضی عملیات کنترل همروندی تراکنشها
- ۶- ضابطه تشخیص کامل بودن زبانهای رابطه‌ای

حساب رابطه‌ای:

دسته دوم از امکانات انجام عملیات در رابطه‌ها, حساب رابطه‌ای است. حساب رابطه‌ای با جبر رابطه‌ای منطقاً معادل است, یعنی برای هر عبارت جبر رابطه‌ای, یک عبارت معادل در حساب رابطه‌ای وجود دارد و برعکس.

تفاوت آنها این است که جبر رابطه‌ای, دستوری است, اما حساب رابطه‌ای توصیفی است.

حساب رابطه‌ای:

■ حساب تاپلی:

در این حساب مفهوم مهمی به نام متغیر تاپلی وجود دارد که تنها مقادیر مجازش، تاپلهای رابطه هستند.

■ حساب میدانی:

در این حساب، متغیر میدانی وجود دارد که از یک میدان مقدار می‌گیرد. در این حساب یک شرط اضافی به نام شرط عضویت وجود دارد.

جلسه دوازدهم

زبان SQL

آنچه در این جلسه می خوانید:

۱- دستورات زبان SQL

۲- انواع داده‌ها

۳- دستورات تعریف داده‌ها

۴- دستورات اعطای مجوز

۵- دستورات پردازش داده‌ها

۶- برخی از توابع در SQL

۷- استفاده از توابع در دستورات

۸- دستورات نوشتن ماژول و رویه

۹- دستورات کنترل جامعیت

۱۰- دستورات کنترل تراکنشها

زبان SQL:

SQL یک زبان رابطه‌ای است و از جمله امکانات مهم آن می‌توان به موارد زیر اشاره کرد:

- دستورات تعریف داده‌ها
- دستورات اعطای مجوز
- دستورات پردازش داده‌ها
- دستورات پردازش داده‌ها به طور ادغام‌شدنی
- دستورات نوشتن ماژول و رویه
- دستورات کنترل جامعیت
- دستورات کنترل تراکنشها

انواع داده‌ها:

- CHARACTER [(length)]
- INTEGER
- DECIMAL [(precision[,scale])]
- SMALLINT
- DOUBLE PRECISION
- REAL
- FLOAT [(precision)]
- NUMERIC [(precision[,scale])]

دستورات تعریف داده‌ها:

■ دستور تعریف جدول:

... نام جدول create table

■ دستور حذف جدول:

نام جدول drop table

دستورات اعطای مجوز:

با این دستورات حق انجام یک یا بیش از یک عمل به کاربر (کاربرانی) داده یا گرفته می‌شود.

▪ دستورات اعطای مجوز:

grant ...

revoke ...

▪ دستورات لغو مجوز:

deny ...

دستورات پردازش داده‌ها:

■ دستور بازیابی

نام صفات `SELECT [ALL | DISTINCT]`

نام جدول `FROM`

[شرطها `WHERE`]

[`GROUP BY column(s)`]

[شرطها `HAVING`]

برخی از توابع در SQL:

- **COUNT:** تعداد مقادیر اسکالر را به دست می‌دهد.
- **SUM:** حاصل جمع مقادیر اسکالر در یک ستون را به دست می‌دهد.
- **AVG:** میانگین مقادیر اسکالر در یک ستون را به دست می‌دهد.
- **MAX:** بیشترین مقدار اسکالر در یک ستون را به دست می‌دهد.
- **MIN:** کمترین مقدار اسکالر در یک ستون را به دست می‌دهد.

مثال استفاده از توابع در دستور بازیابی:

```
SELECT MIN(GRADE) , MAX(GRADE) FROM STUDENT  
WHERE TR='2' AND YRYR='78-79' AND COID='COM222';
```

بالاترین و پایین ترین نمره در درس COM222 در ترم دوم ۸۰-۷۹ را از جدول STUDENT بازیابی می کند.

دستورات پردازش داده‌ها:

■ دستور درج:

INSERT INTO نام جدول Values(مقادیر);

■ دستور حذف:

DELETE FROM نام جدول [WHERE شرطها]

■ دستور بروزرسانی:

UPDATE نام جدول SET تغییرات [WHERE شرطها]

دستورات نوشتن ماژول و رویه:

■ دستور نوشتن رویه (زیربرنامه):

... نام رویه create procedure

■ دستور نوشتن تابع:

... نام تابع create function

دستورات کنترل جامعیت:

▪ دستورات اعلان محدودیت میدانی و صفتی:

... نوع داده نام دامنه `create domain`

▪ دستور مکانیسم اظهار:

شرطها `check` نام قید `create assertion`

▪ دستور مکانیسم رهانا:

... نام `create trigger`

دستورات کنترل تراکنشها:

برای کنترل جنبه‌هایی از تراکنش دستورهایی زیر وجود دارند:

set transaction

commit

rollback

جلسه سیزدهم

فرمالترسازی رابطه‌ها

آنچه در این جلسه می خوانید:

۱- نرمالترسازی رابطه‌ها

۶- وابستگی تابعی نامهم (بدیهی)

۲- آنومالی در عملیات ذخیره‌سازی

۷- وابستگی تابعی چندمقداری

۳- صورتهای (فرمهای) نرمال

۸- روابط 1NF, 2NF, 3NF, BCNF, 5NF

۴- وابستگی تابعی

۹- مزایای نرمالترسازی

۵- وابستگی تابعی تام (کامل)

۱۰- معایب نرمالترسازی

چرا نرمالترسازی رابطه‌ها:

یک رابطه، هرچند نرمال، بازهم ممکن است در عملیات ذخیره‌سازی (درج، حذف و بهنگام‌سازی) مشکلاتی داشته باشد.

رابطه نرمال: رابطه‌ای است که تمام صفات آن تک‌مقداری باشند

چرا نرمالتر سازی رابطه‌ها:

نمره	مقطع	نام دانشجو	شماره درس	شماره دانشجویی
۱۰	لیسانس	St1	Com111	۷۸۰۱۰۲۲۲
۷	لیسانس	St1	Mat120	۷۸۰۱۰۲۲۲
۱۲	لیسانس	St1	Art302	۷۸۰۱۰۲۲۲
۱۳	فوق لیسانس	St2	Com111	۷۷۰۲۰۱۱۱
۹	فوق لیسانس	St2	Com115	۷۷۰۲۰۱۱۱
?	لیسانس	St3	Com111	۷۷۰۱۰۳۳۳
۱۶	فوق لیسانس	St4	Mat122	۷۸۰۲۰۴۴۴
۱۱	فوق لیسانس	st4	phy444	۷۸۰۲۰۴۴۴

<78020777,st7,لیسانس>
 <7701033,com111,?>
 <78010222>

در عمل درج:
 در عمل حذف:
 در بهنگام سازی:

آنومالی

آنومالی (وضع غیرعادی) در عملیات ذخیره‌سازی:

- عدم امکان انجام یک عمل (که منطقاً باید قابل انجام باشد)

- بروز پیامد بد پس از انجام یک عمل

- بروز فزونکاری در سیستم در انجام یک عمل

صورت‌های (فرم‌های) نرمال:

- صورت نرمال نخست (1NF: First Normal Form)
- صورت نرمال دوم (2NF: Second Normal Form)
- صورت نرمال سوم (3NF: Third Normal Form)
- صورت نرمال بایس-کاد (BCNF: Boyce-Codd Normal Form)
- صورت نرمال چهارم (4NF: Fourth Normal Form)
- صورت نرمال پنجم (5NF: Fifth Normal Form)
- صورت نرمال ششم (6NF: Sixth Normal Form)
- صورت نرمال میدان-کلیدی (DKNF: Domain-Key Normal Form)
- صورت نرمال تحدید-اجتماع (RUNF: Restriction-Union Normal Form)

رابطه 1NF:

رابطه‌ای 1NF است که هر صفت خاصه آن در هر تاپل، تک‌مقداری باشد، به بیان دیگر، صفت چندمقداری نداشته باشد.

وابستگی تابعی:

فرض کنید که R یک متغیر رابطه‌ای و A و B دو زیرمجموعه دلخواه از عنوان R باشند. می‌گوییم B با A وابستگی تابعی دارد و چنین نمایش می‌دهیم:

$$A \rightarrow B$$

اگر و فقط اگر در هر میدان ممکن از متغیر رابطه‌ای R ، به هر مقدار A فقط یک مقدار B متناظر باشد.

Student (stID, stName, stLevel, stMajoring)

$$\left\{ \begin{array}{l} \text{stID} \rightarrow \text{stName} \\ \text{stID} \rightarrow \text{stLevel} \\ \text{stID} \rightarrow \text{stMajoring} \end{array} \right\}$$

وابستگی تابعی تام (کامل):

اگر X و Y دو زیرمجموعه از مجموعه عنوان رابطه R باشند، می‌گوییم Y با X وابستگی تابعی تام دارد و چنین نشان می‌دهیم: $X \Rightarrow Y$

اگر و فقط اگر Y با X وابستگی داشته باشد. ولی با هیچ زیرمجموعه‌ای از X وابستگی تابعی نداشته باشد.

رابطه 2NF:

رابطه‌ای 2NF است که اولاً 1NF باشد و ثانیاً تمام صفات غیرکلید با کلید اصلی وابستگی تابعی تام داشته باشند. به عبارت دیگر هر صفت غیرکلید با کلید اصلی بطور کاهش‌ناپذیر وابسته باشد.

رابطه 3NF:

رابطه‌ای 3NF است که 2NF باشد و هر صفت غیرکلید با کلید اصلی، وابستگی تابعی بی‌واسطه داشته باشد.

صورت نرمال بایسی – کاد (BCNF):

▪ تعریف اول:

رابطه‌ای BCNF است که در آن هر دترمینان، کلید کاندید باشد. این تعریف ساده شده و غیر صوری است.

یادآوری:

در رابطه $R(A, B, \dots)$ اگر $A \rightarrow B$ باشد، می‌گوییم A دترمینان است.

▪ تعریف دوم

▪ تعریف سوم

▪ تعریف چهارم

وابستگی تابعی نامهم (بدیهی):

اگر در $R(A, B, C, \dots)$ داشته باشیم: $B \subseteq A$ ، در این صورت: $A \rightarrow B$ یک وابستگی تابعی نامهم است.

به بیان دیگر اگر B زیرمجموعه‌ای از A باشد، در این صورت $A \rightarrow B$ یک وابستگی نامهم است.

وابستگی تابعی چندمقداری:

• تعریف اول:

در رابطه $R(X, Y, Z)$ با صفات ساده یا مرکب X ، Y و Z می‌گوییم که Y با X وابستگی تابعی چندمقداری دارد و چنین نمایش می‌دهیم: $X \rightarrow Y$
اگر به یک مقدار X ، مجموعه‌ای از مقادیر Y متناظر باشند.

▪ تعریف دوم

رابطه 4NF:

رابطه‌ای 4NF است که BCNF باشد و در آن وابستگی تابعی چند مقداری مهم وجود نداشته باشد.

رابطه 5NF:

رابطه‌ای 5NF است که تمام وابستگی‌های پیوندی آن ناشی از کلیدهای کاندید آن باشد. در نتیجه:

اگر بتوانیم یک وابستگی پیوندی در رابطه R پیدا کنیم که در همه پرتوهایش کلید کاندید R وجود نداشته باشد، رابطه 5NF نیست.

مزایای نرمالتر سازی:

- ارائه یک طراحی بهتر و واضح تر با کمترین اختلاط اطلاعات
- کاهش بعضی انواع افزونگی
- کاهش بعضی آنومالیها
- تسهیل اعمال بعضی قواعد جامعیت

معایب نرمالترسازی:

- بروز فزونکاری در سیستم در عمل بازیابی
- ایجاد نوعی افزونگی از نوع افزونگی در سطح ادراکی
- زمانگیر بودن فرآیند نرمالترسازی به ویژه اگر محیط عملیاتی بزرگ و تعداد رابطه‌ها زیاد باشد.
- تصمیم‌گیری دشوار در اثر تعدد تجزیه‌ها
- ممکن وابستگی بین مجموعه صفات یک خردجهان، وابستگی‌های به جز وابستگی تابعی و پیوندی باشد و بنابراین سبب مطرح شدن قواعد جامعیت دیگر و نیز ضوابط دیگری برای تجزیه رابطه‌ها شود و ...

پایان