

COMP 3380 Winter 2023 Project

Part 1: Designing a Database

Group 16

Christie, Nikolaas

christ37@myumanitoba.ca

Duong, Edwin

duonge1@myumanitoba.ca

Zhang, Linpu

zhangl53@myumanitoba.ca

Summary of Dataset

The Chinook sample database represents a digital media store, namely the media it would sell, its customers, its employees, and sample invoice information that would represent transactions between the customers and the store.

- Media-related data was created using real data from the Apple iTunes library.
- Customer and employee information was created using fictitious names and addresses that can be located on Google maps and other well formatted data (phone, fax, email, etc.)
- Invoice/Sales information was auto generated using random data for a four-year period.

In total, the Chinook sample database includes:

- Eleven tables.
- A variety of indexes, primary and foreign key constraints.
- Over 15,000 rows of data.

Figures

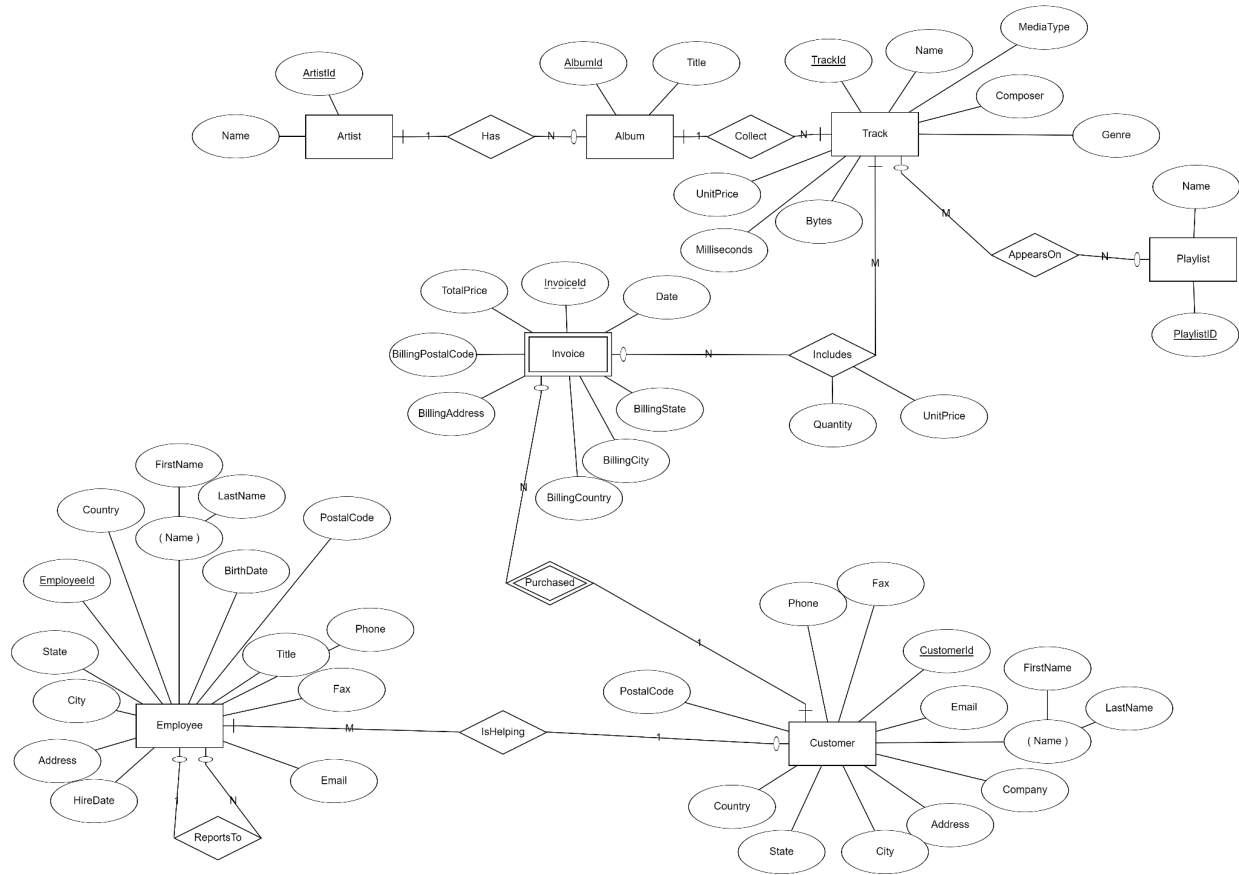


Figure 1. ER Model

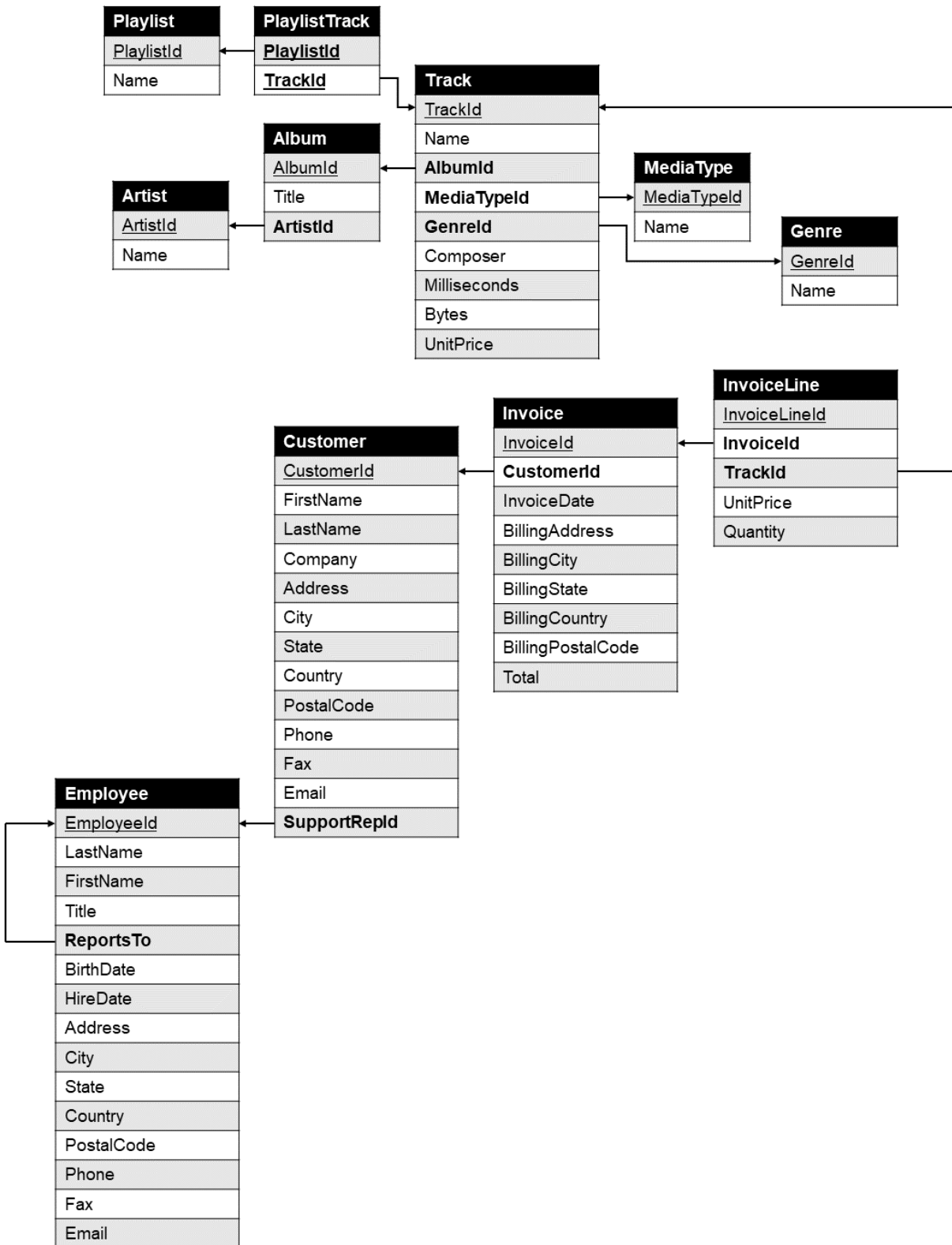


Figure 2. Relational Model

Justification of Participation and Cardinality Constraints

ReportsTo

Some people don't have to report to anyone. For example, the general manager is the boss so there is no need to report. The employee in this relationship has optional participation.

Not every employee is a manager and has subordinates. The manager in this relationship has optional participation.

An employee only reports to one manager. A manager can manage many employees. So the manager-to-employee cardinality ratio is 1:N, or one-to-many.

IsHelping

Every customer needs to be supported by a sales support agent. This is common in companies to achieve better customer satisfaction. The customer in this relationship has total participation.

Some employees such as managers and IT staff don't directly connect to customers. The employee in this relationship has optional participation.

Every customer is only supported by one sales support agent and a sales support can provide service to many customers. So the customer-to-employee cardinality ratio is M:1, or many-to-one.

Purchased

Every invoice must have a customer because every purchase is made by a certain customer. Invoice in this relationship has total participation.

Not every customer has an invoice (already signed up but has not made a purchase). The customer in this relationship has optional participation.

A customer can make many invoices and an invoice can only be made by one customer. So the customer-to-invoice cardinality ratio is 1:N.

Includes

You must have at least one item to buy for every order. Invoice in this relationship has total participation.

A track may not be purchased by anyone. Track in this relationship has optional participation.

An invoice can include many tracks and a track can be in many invoices. So the track-to-invoice cardinality ratio is M:N.

AppearsOn

It is reasonable to create an empty playlist and add tracks later. So the playlist can be empty. Playlist in this relationship has optional participation.

Some tracks may not be included in any playlist. Track in this relationship has optional participation.

A playlist can have many tracks and a track can be in many playlists. Playlist-to-track cardinality ratio is M:N.

Has

An artist may not have published any album yet. Artist in this relationship has optional participation.

An album must be made by a certain artist. Album in this relationship has total participation.

An artist can release many albums and an album can only be made by one artist. Album-to-artist cardinality ratio is M:1.

Collect

An album must have at least one song. Album in this relationship is total participation.

A track must be in an album. Even when an artist releases a single song, it's still in a one-song album which this song is included in. Track in this relationship is total participation.

An album can have many tracks. A track can only be included in one album. In a real music app, even though one song can be in different albums (like the original release and a Greatest

Hits album), it is treated as multiple different songs but in the same name in practice because of convenience of management and copyright issues. So the album to track cardinality is 1:N.

Translating ER diagram to relational model:

Strong entity **Employee** is represented by a relation. All attributes are preserved.

1:N relationship **ReportsTo** is translated into an attribute of Employee which is a foreign key referencing the primary key of Employee.

Strong entity **Customer** is represented by a relation. All attributes are preserved.

1:N relationship **IsHelping** is translated into an attribute SupportRepId of Customer which is a foreign key referencing primary key of Employee.

Weak entity **Invoice** and 1:N relationship **Purchased** is represented by a relation. All attributes are preserved and the primary key of identifying entity Customer is included. Primary key for the relation Invoice is the surrogate key InvoiceId.

M:N relationship **Includes** is translated into a 3rd relation called InvoiceLine. All attributes of the relationship are preserved and primary keys of Invoice and Track are included. We created a surrogate key InvoiceLineId as the primary key of the relation. Notice the attribute **UnitPrice** might be different from the one in Track because this is the price when the order is made and may be influenced by a discount, etc.

Strong entities **Track**, **Album**, **Artist** and **Playlist** are represented by relations. All attributes are preserved.

1:N relationship **Has** and **Collect** are translated to foreign keys of Album and Track.

Genre and **MediaType** are attributes that only contain common values and repeatedly appear on each row of Track. So translating them into two relations includes original values and a surrogate key. Include foreign keys GenreId and MediaTypeId in Track referencing to primary keys of the two relations.

M:N relationship **AppearsOn** is translated into a 3rd relation called PlaylistTrack. Primary keys of Playlist and Track are included as foreign keys and the combination of them is the primary key of the relation.

The relation mode is already normalized after translation.

Appendix

Source data was obtained from the Chinook sample database using the Microsoft Public License (Ms-PL).

<https://github.com/w3c/csvw/tree/gh-pages/examples/tests/scenarios/chinook>