



ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA

EDUCACIÓN
PROFESIONAL

Buen diseño de bases de datos I

Gestión de Datos

Vicente Calisto

Educación Profesional - Escuela de Ingeniería

Clase diseñada por Matías Toro para GDD, DCDPP 2022

El uso de apuntes de clases estará reservado para finalidades académicas. La reproducción total o parcial de los mismos por cualquier medio, así como su difusión y distribución a terceras personas no está permitida, salvo con autorización del autor.

Hasta ahora

Conocemos el modelo relacional y SQL por lo que podemos comenzar a diseñar una base de datos, pero... ¿Sabemos si lo estamos haciendo bien?

Un error en la modelación puede ser muy costoso!

Por ejemplo: olvidar añadir una columna



Errores en los datos

Investigaciones precisas necesitan valores precisos, sin embargo, tener datos libres de errores es prácticamente imposible.

La propagación de errores tanto a nivel numérico, como computacional y humano debe ser correctamente manejado para obtener resultados confiables

$$\text{PRECISE NUMBER} + \text{PRECISE NUMBER} = \text{SLIGHTLY LESS PRECISE NUMBER}$$

$$\text{PRECISE NUMBER} \times \text{PRECISE NUMBER} = \text{SLIGHTLY LESS PRECISE NUMBER}$$

$$\text{PRECISE NUMBER} + \text{GARBAGE} = \text{GARBAGE}$$

$$\text{PRECISE NUMBER} \times \text{GARBAGE} = \text{GARBAGE}$$

$$\sqrt{\text{GARBAGE}} = \text{LESS BAD GARBAGE}$$

$$(\text{GARBAGE})^2 = \text{WORSE GARBAGE}$$

$$\frac{1}{N} \sum (\text{N PIECES OF STATISTICALLY INDEPENDENT GARBAGE}) = \text{BETTER GARBAGE}$$

$$(\text{PRECISE NUMBER})^{\text{GARBAGE}} = \text{MUCH WORSE GARBAGE}$$

$$\text{GARBAGE} - \text{GARBAGE} = \text{MUCH WORSE GARBAGE}$$

$$\frac{\text{PRECISE NUMBER}}{\text{GARBAGE} - \text{GARBAGE}} = \text{MUCH WORSE GARBAGE, POSSIBLE DIVISION BY ZERO}$$

$$\text{GARBAGE} \times 0 = \text{PRECISE NUMBER}$$



¿Como diagnosticar la falta de calidad?

Para diagnosticar la calidad de un conjunto de datos podemos evaluar las diferentes dimensiones en ellos:

- **Exactitud:** Contrastar valor real con valor en DB.
- **Validez:** Contrastar un conjunto de reglas de la definición de los datos en la DB.
- **Completitud:** Contar el porcentaje de datos faltantes.
- **Consistencia:** Distintas fuentes de datos hacen sentido entre sí.
- **Uniformidad:** Los datos se encuentran en la misma escala.
- Etc...



Diseño de base de datos

Análisis de requisitos

Usuarios



Requisitos



Diseño conceptual de bases de datos

Requisitos

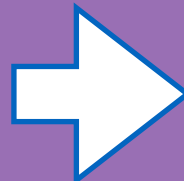


Modelo entidad-relación



Diseño lógico de bases de datos

Modelo entidad-relación



Modelo relacional



Diseñando una base de datos

- En la práctica es imposible saber de antemano todos los requisitos que debe cumplir una bd.
- De la mano con eso, el esquema de la base de datos va cambiando en el tiempo.
- Un esquema bien diseñado no solo nos permite consultar con facilidad y guardar los datos de forma óptima, si no que también permite modificarlo y aumentarlo con menos dolores de cabeza.
- **Los errores en el diseño son muy costosos a la larga!**



Diseño conceptual de la BD

- ¿Por qué diseñar y diagramar la base de datos?:
 - Identificar las entidades.
 - Entender cómo se asocian esas entidades.
 - Visualizar las restricciones del dominio.
 - Para lograr un buen diseño!
 - Para mantener el esquema bien documentado.



Diagramas E/R



Objetos básicos

Entidad

Producto

Atributo

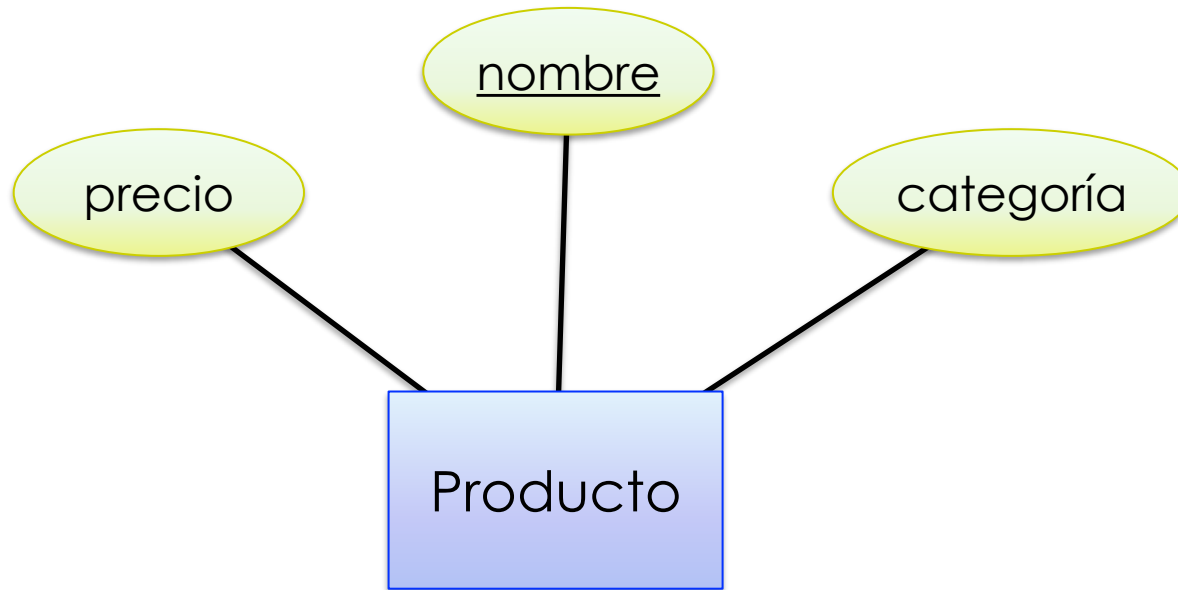
nombre

Relación

compra



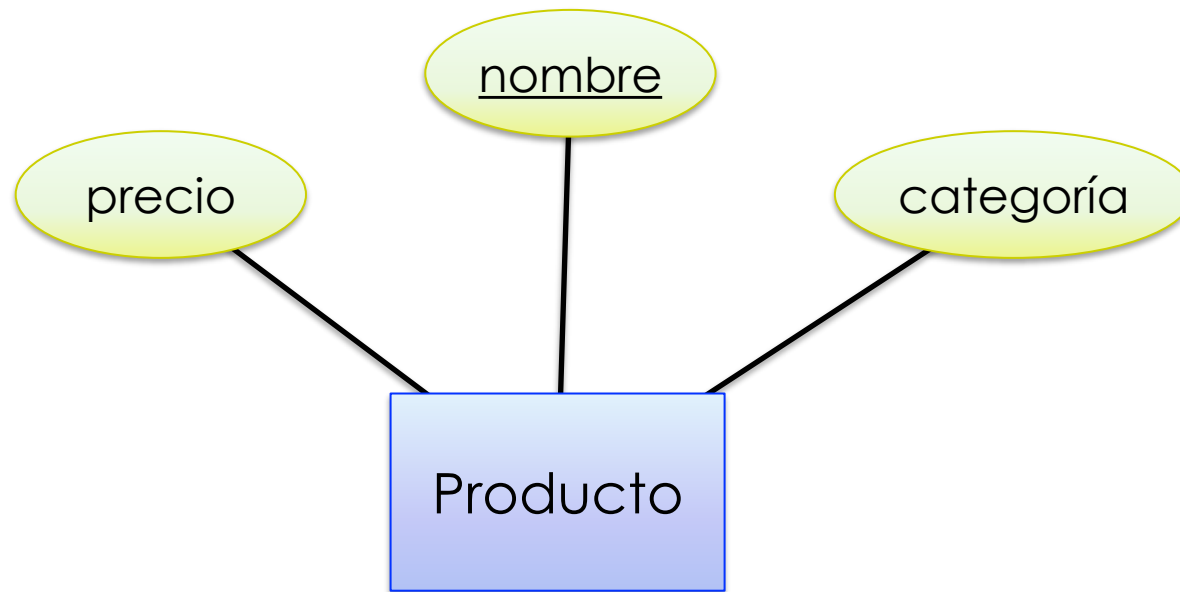
Diagramas E/R: entidad con sus atributos



Obligatorio: cada entidad debe tener una **llave**,
i.e. un conjunto de **atributos** *mínimo*
cuyos valores identifican de manera unívoca
a cada **entidad** del conjunto



Diagramas E/R: entidad con sus atributos

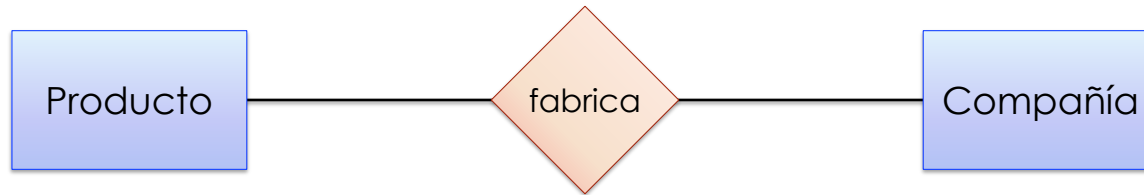


Intuitivamente:

| <u>nombre</u> | precio | categoría |
|------------------|--------|-----------|
| Cusqueña | 1000 | Cerveza |
| Stella Artois | 745 | Cerveza |
| Santa Ema Merlot | 6790 | Vino |
| Corona | 1200 | Cerveza |

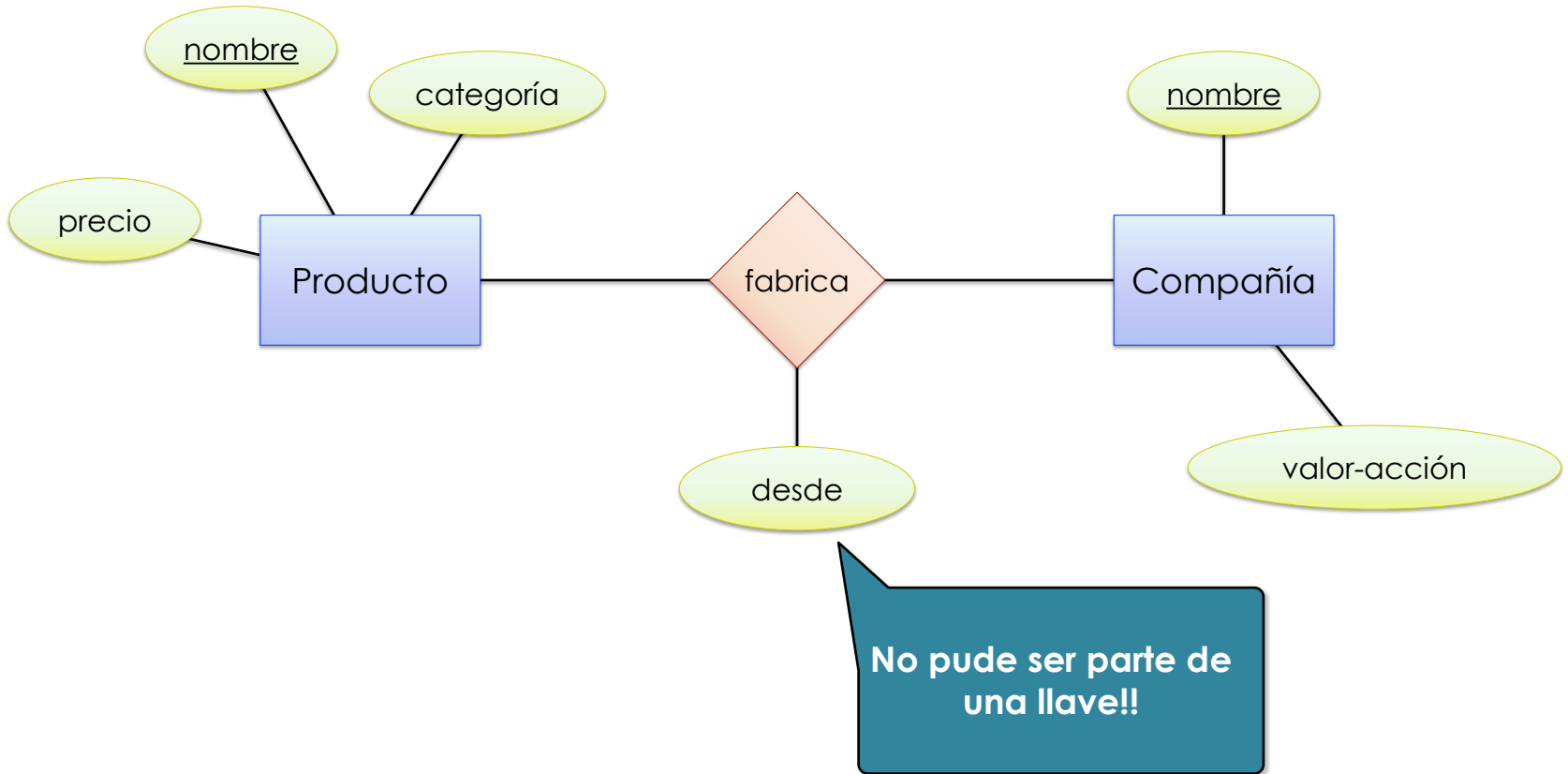


Dos entidades relacionadas

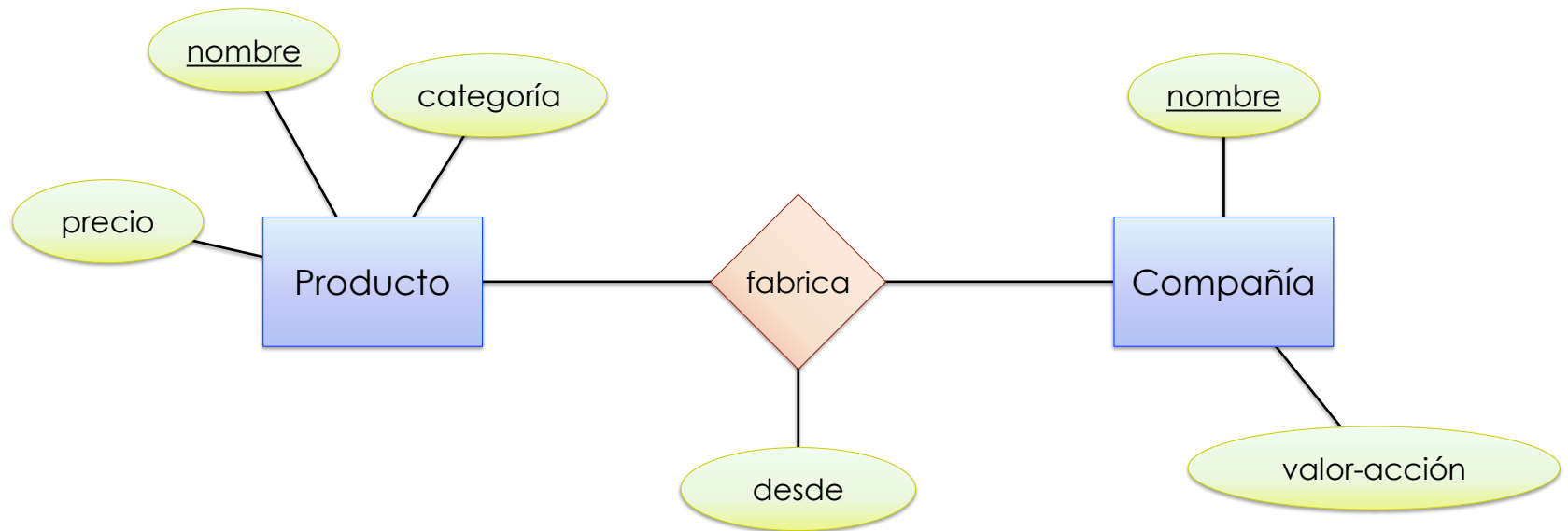


Relaciones Binarias

Entidad con sus atributos



Relaciones Binarias



Intuitivamente (spoilers):

Producto

| <u>nombre</u> | precio | categoría |
|------------------|--------|-----------|
| Cusqueña | 1000 | Cerveza |
| Stella Artois | 745 | Cerveza |
| Santa Ema Merlot | 6790 | Vino |
| Corona | 1200 | Cerveza |

fabrica

| <u>p_nombre</u> | <u>p_compañía</u> | desde |
|------------------|-------------------|-------|
| Cusqueña | Cusqueña | 1908 |
| Stella Artois | AB InBev | 1926 |
| Santa Ema Merlot | Viña Santa Ema | 1956 |
| Corona | AB InBev | 1926 |

Compañía

| <u>nombre</u> | valor-acción |
|----------------|--------------|
| Viña Santa Ema | 5500 |
| Ab InBev | 1224 |
| Cusqueña | 5678 |



Relaciones Binarias

Multiplicidades

n a n

0 o más



Una compañía puede fabricar varios productos

n a 0 o 1



Un producto se fabrica como máximo por una compañía

0 o 1 a n



Un producto es fabricado ser fabricado por muchas compañías

Una compañía fabrica como máximo un producto 🤔

0 o 1 a 0 o 1



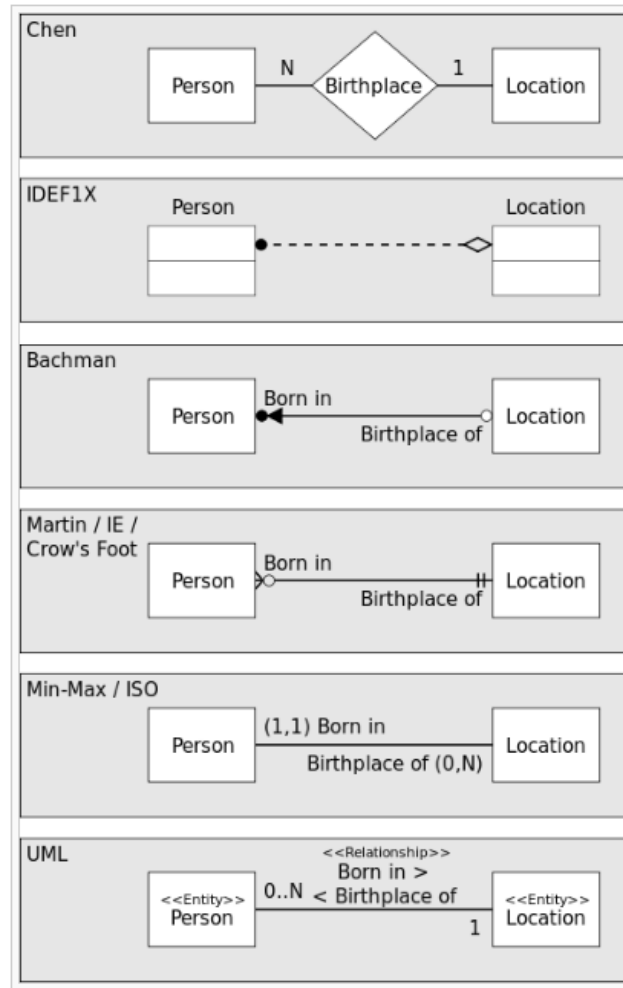
Un producto se fabrica como máximo por una compañía

Una compañía fabrica como máximo un producto 🤔



Relaciones Binarias

Multiplicidades

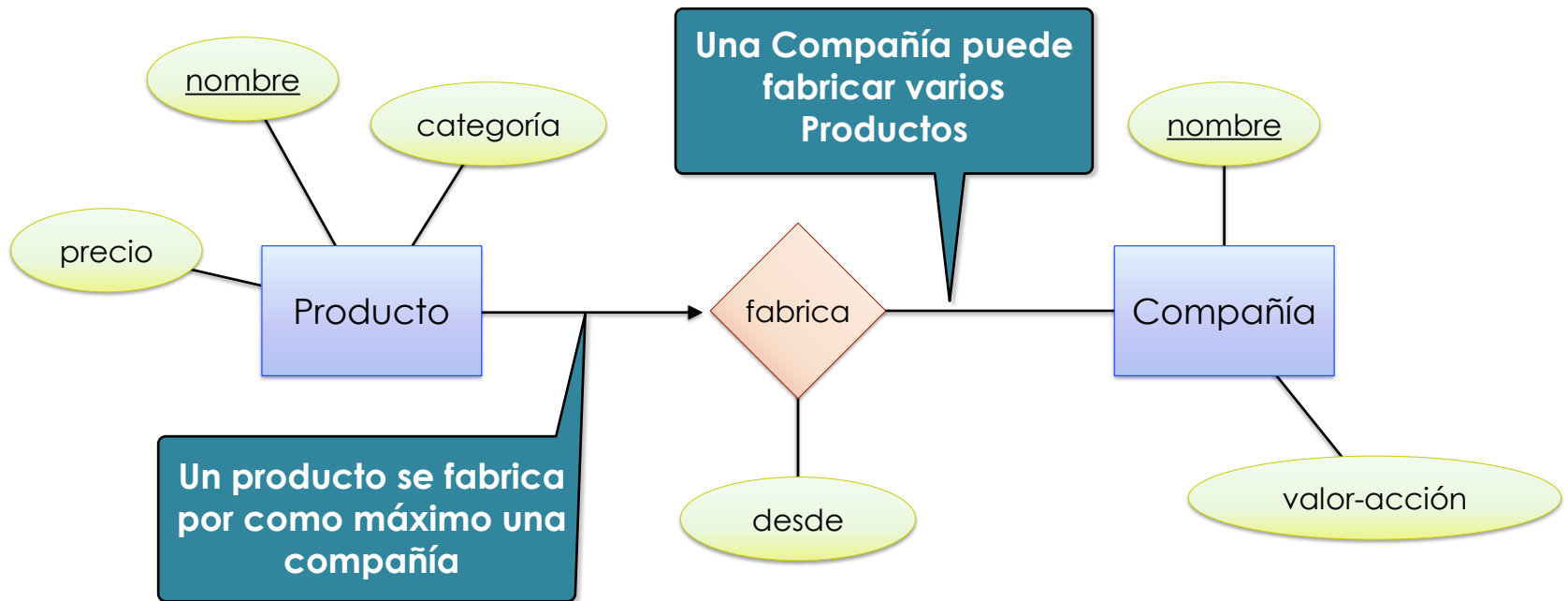


Muchas convenciones distintas



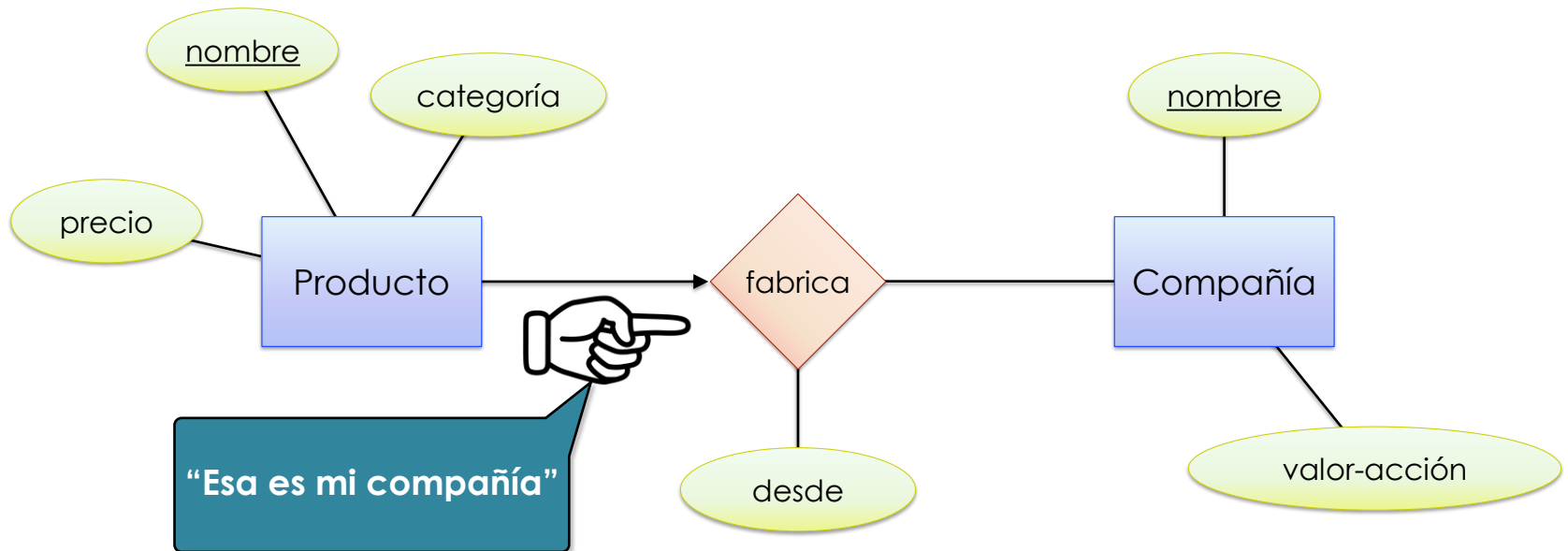
Relaciones Binarias

Sólo utilizaremos esta convención:



Relaciones Binarias

Sólo utilizaremos esta convención:



Diagramas E/R

Relaciones Múltiples



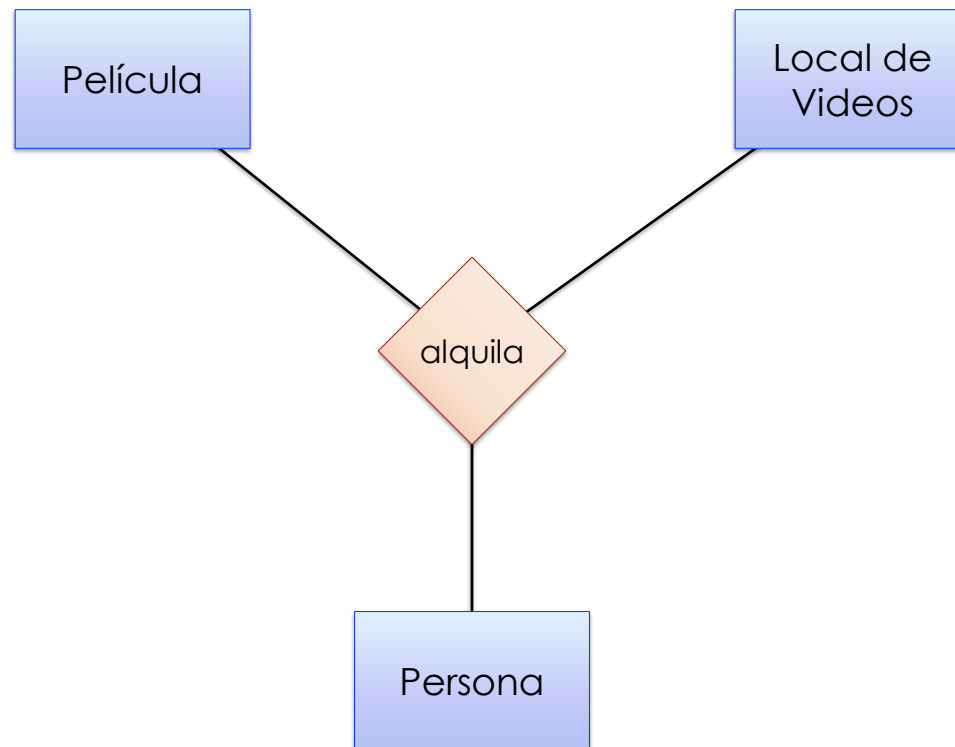
Relaciones Múltiples

¿Cómo se puede modelar un alquiler que involucra **Personas**, **Películas** y **Locales de Videos**?



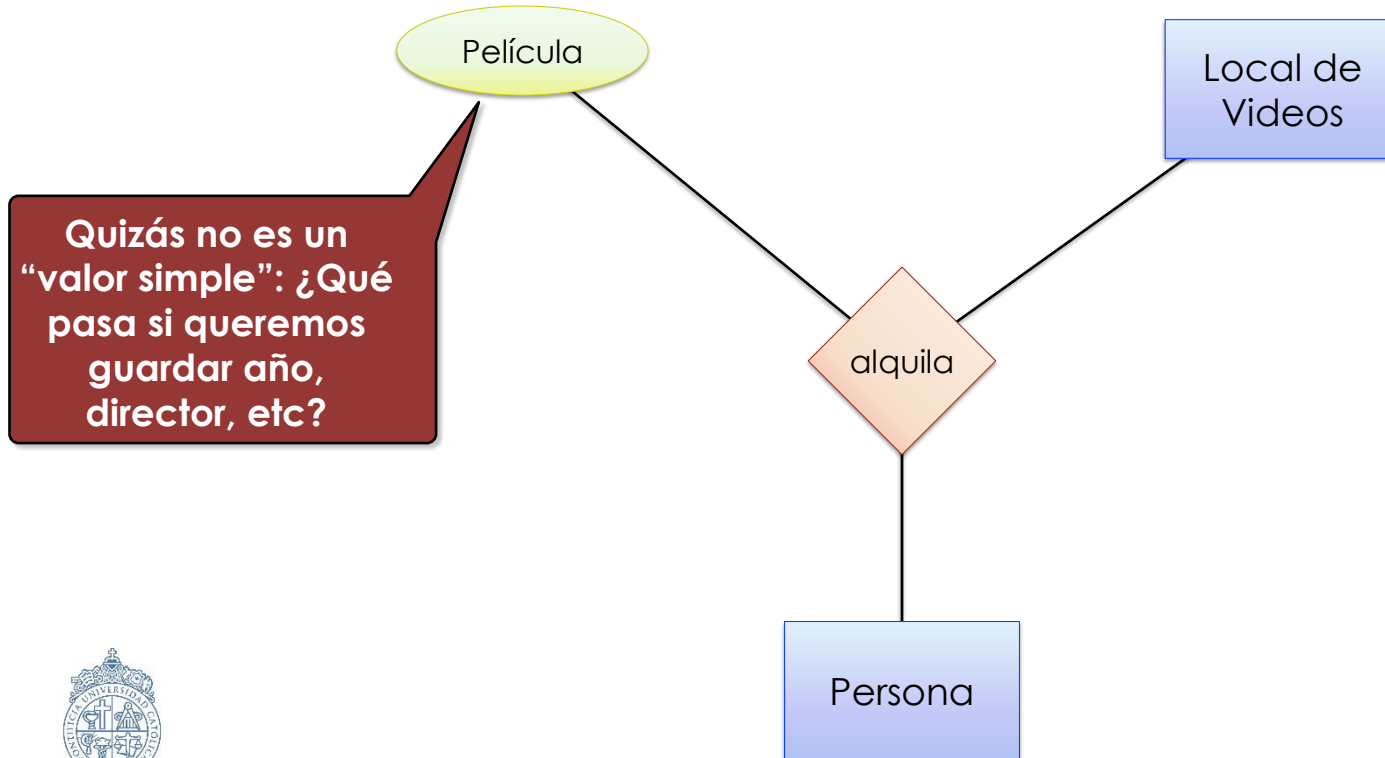
Relaciones Múltiples

¿Cómo se puede modelar un alquiler que involucra **Personas**, **Películas** y **Locales de Videos**?



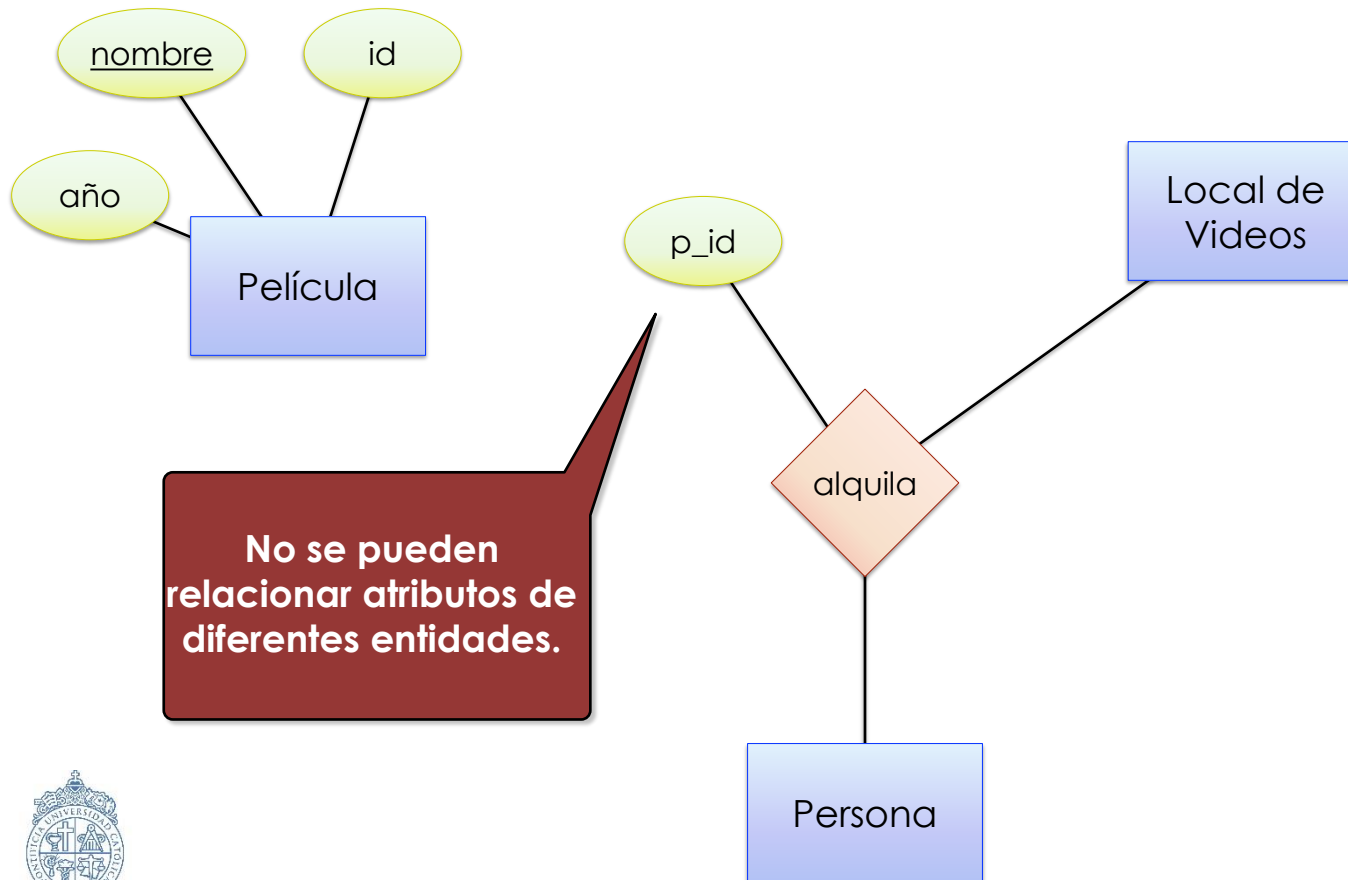
Relaciones Múltiples

¿Por qué no un atributo?



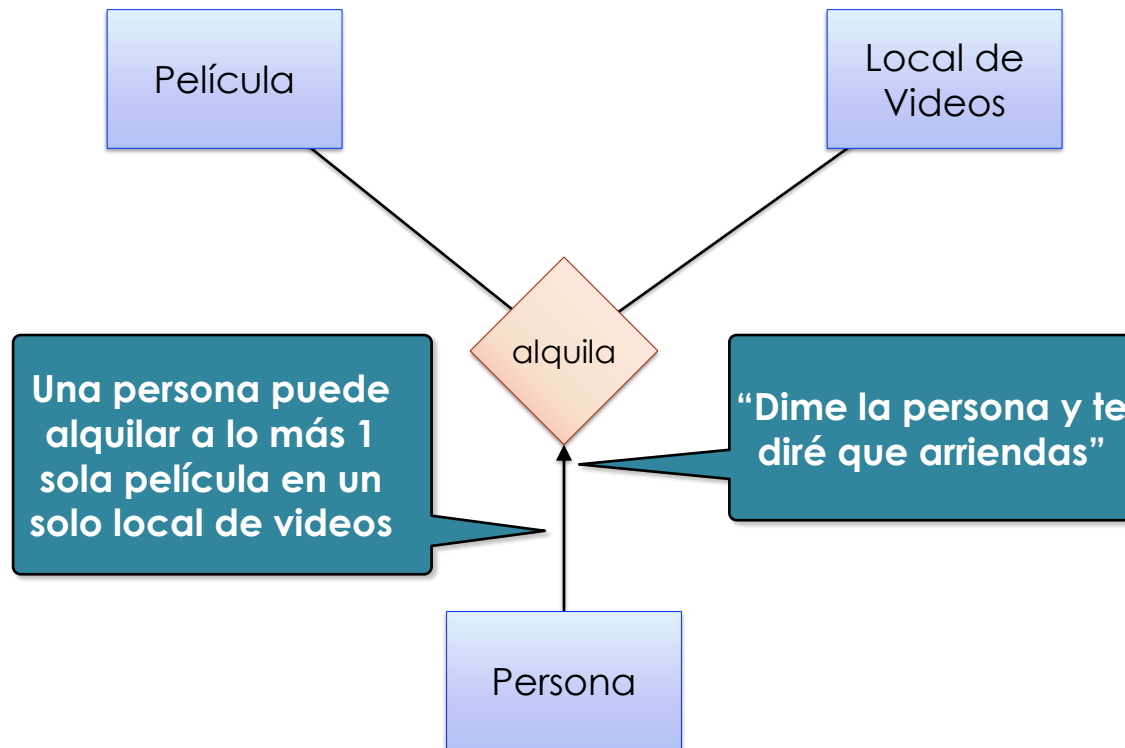
Relaciones Múltiples

¿Y ahora?



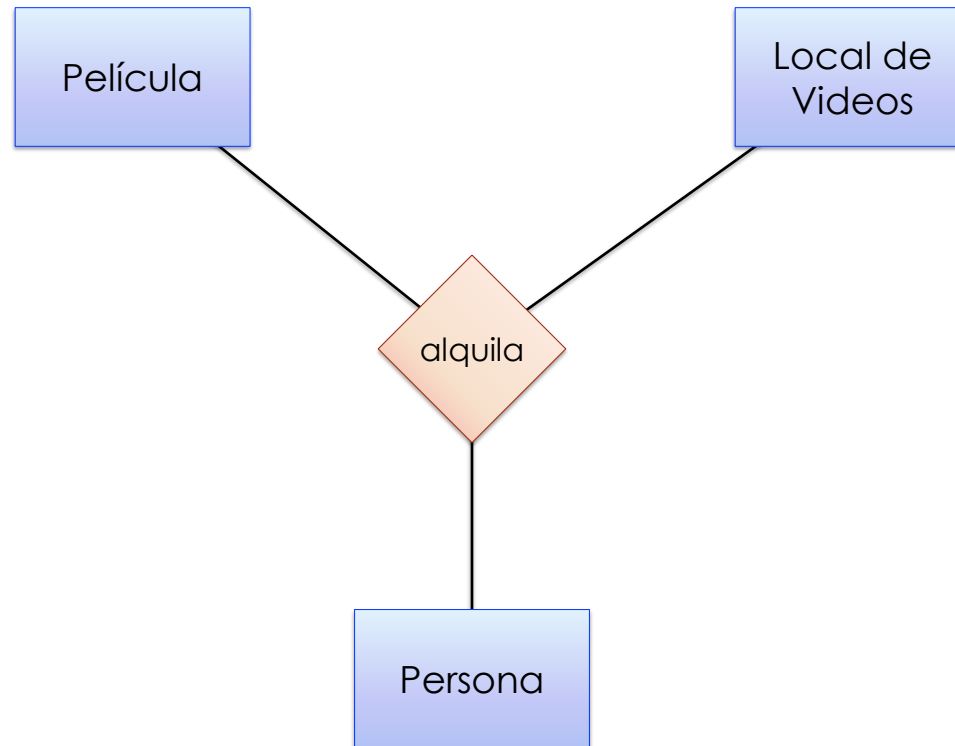
Relaciones Múltiples

¿Qué significa esto?



Relaciones Múltiples

¿Y qué pasa si decimos que una **persona** puede alquilar **varias películas** pero de **un solo local de video**?

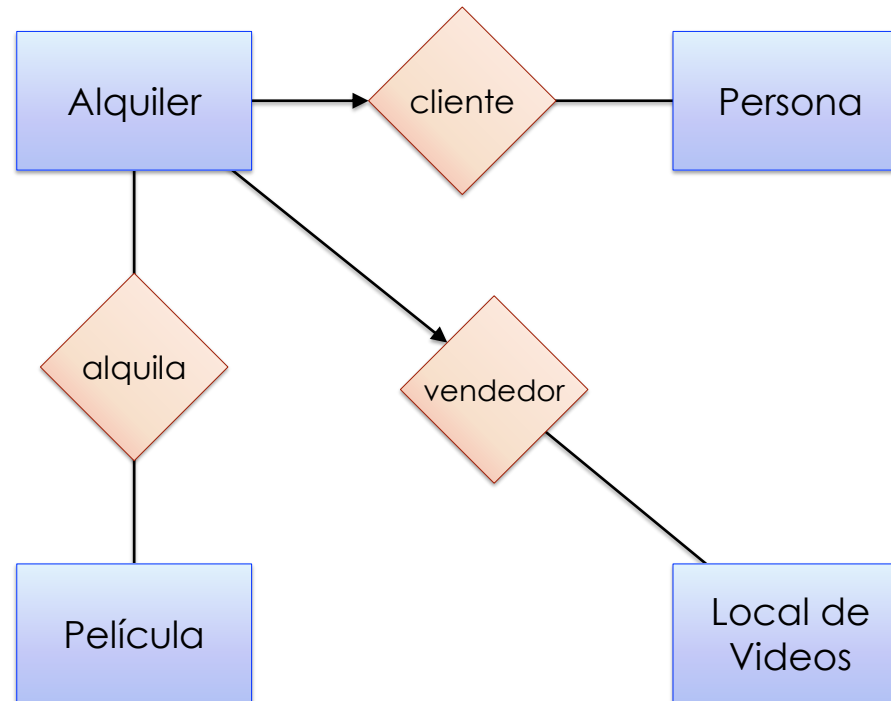


Ups...



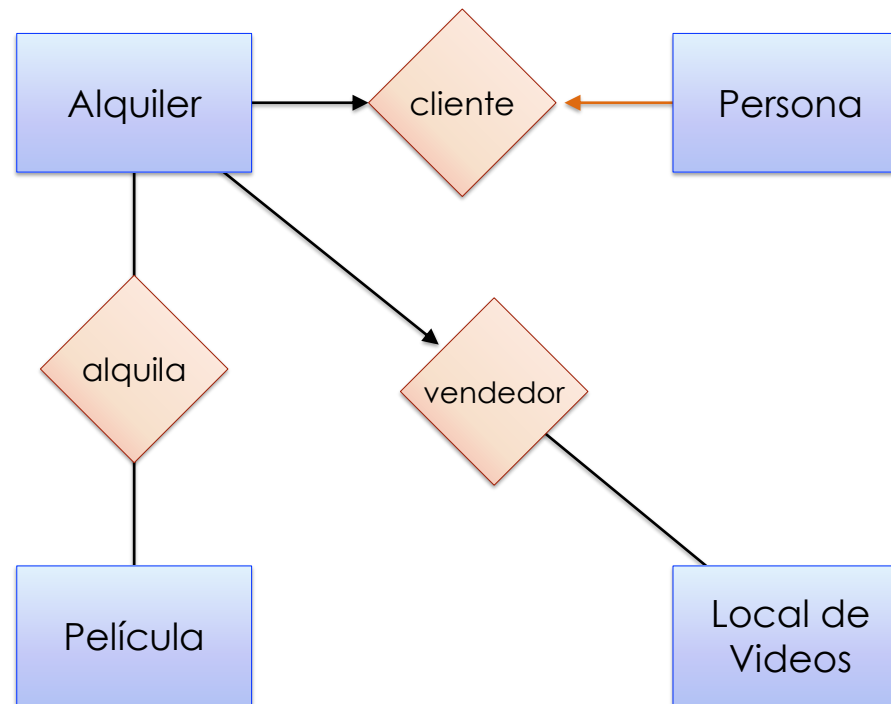
Relaciones Múltiples

¿Qué pasa si usamos solo relaciones binarias?



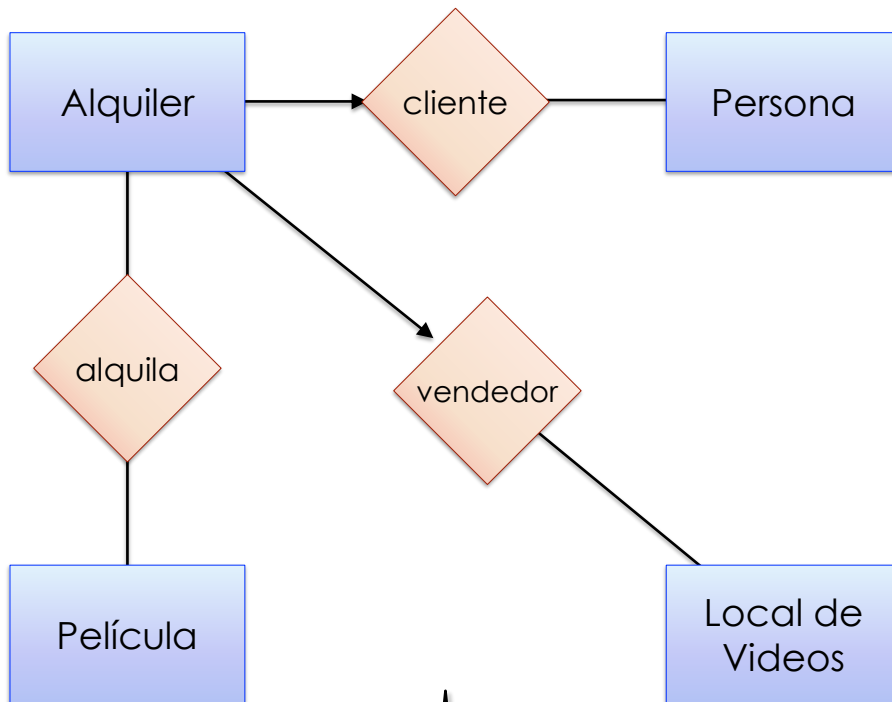
Relaciones Múltiples

¿Y qué pasa si decimos que una **persona** puede alquilar **varias películas** pero de **un solo local de video**?

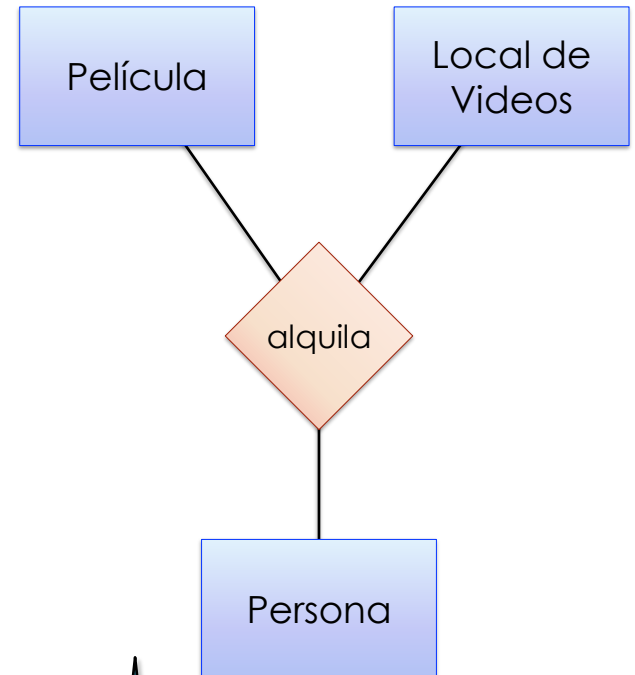


Relaciones Múltiples

¿Cuál es mejor?



Más flexible

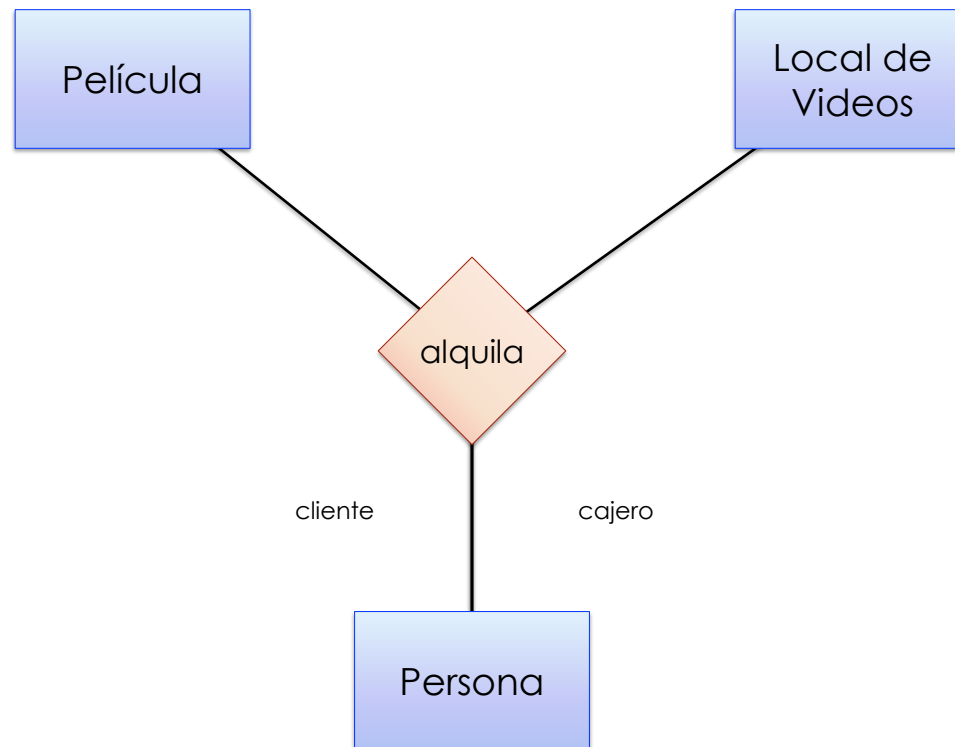


Más conciso



Relaciones Múltiples

Una entidad puede participar más de una vez en una relación

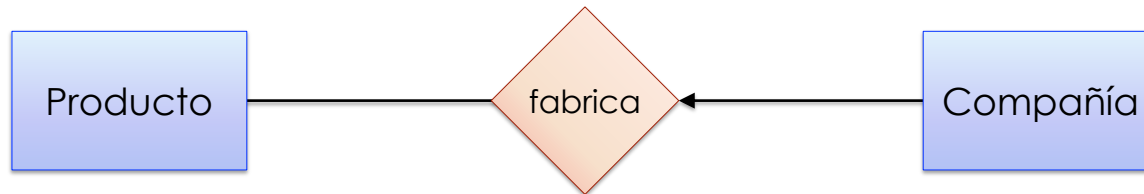
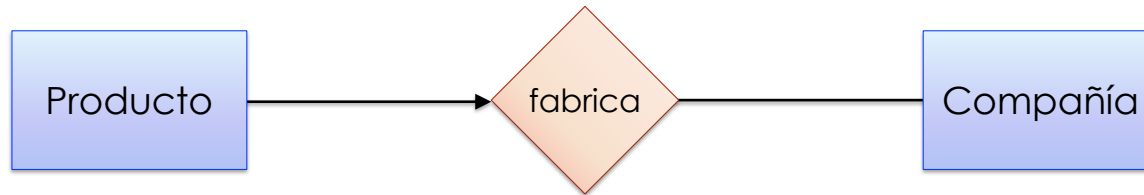


Restricciones Avanzadas



Restricciones Avanzadas

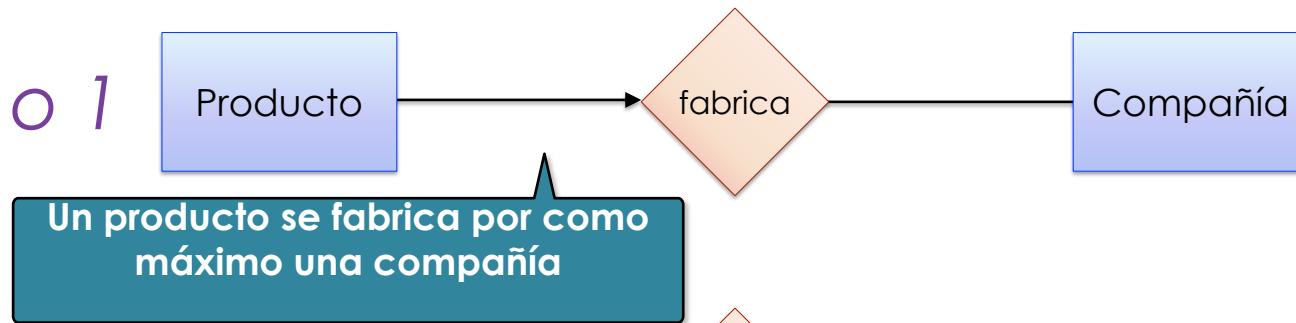
¿Que es más natural?



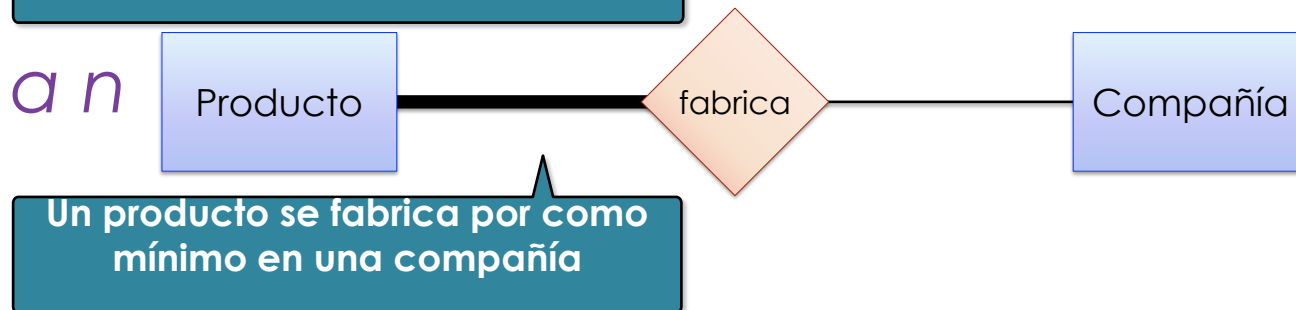
Restricciones Avanzadas

Participación y participación total

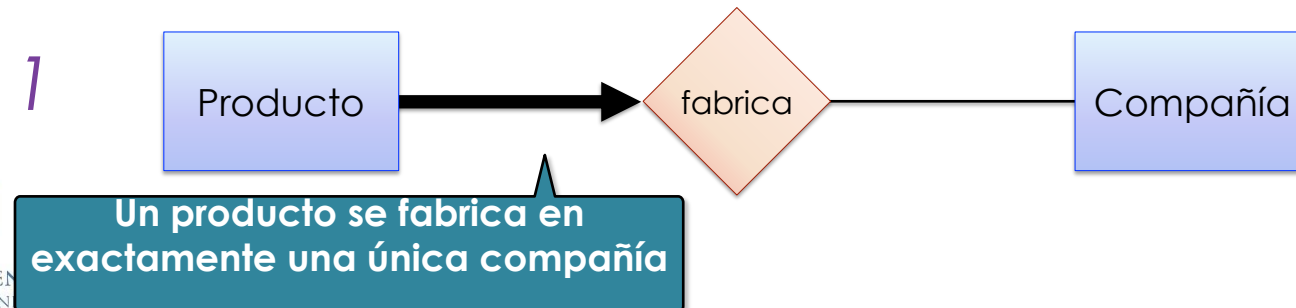
na 0 o 1



na 1 a n



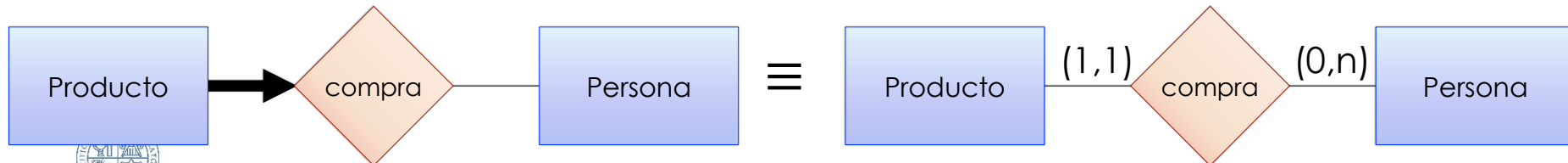
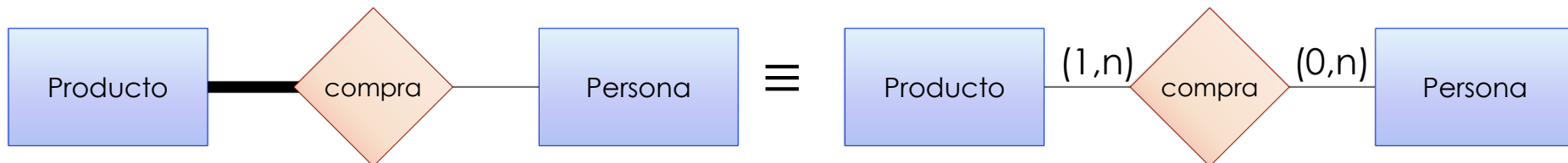
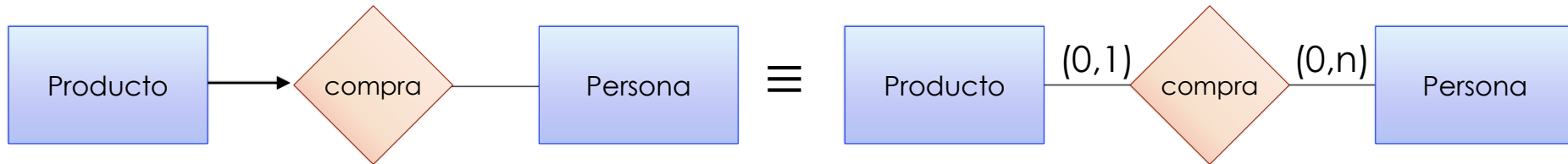
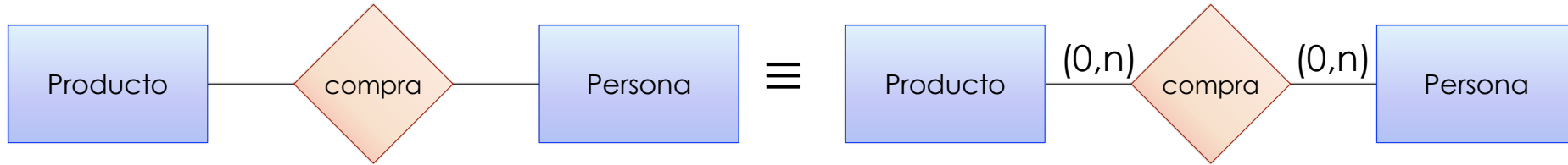
na 1



Equivalencia de multiplicidades

Ramakrishnan

min-max

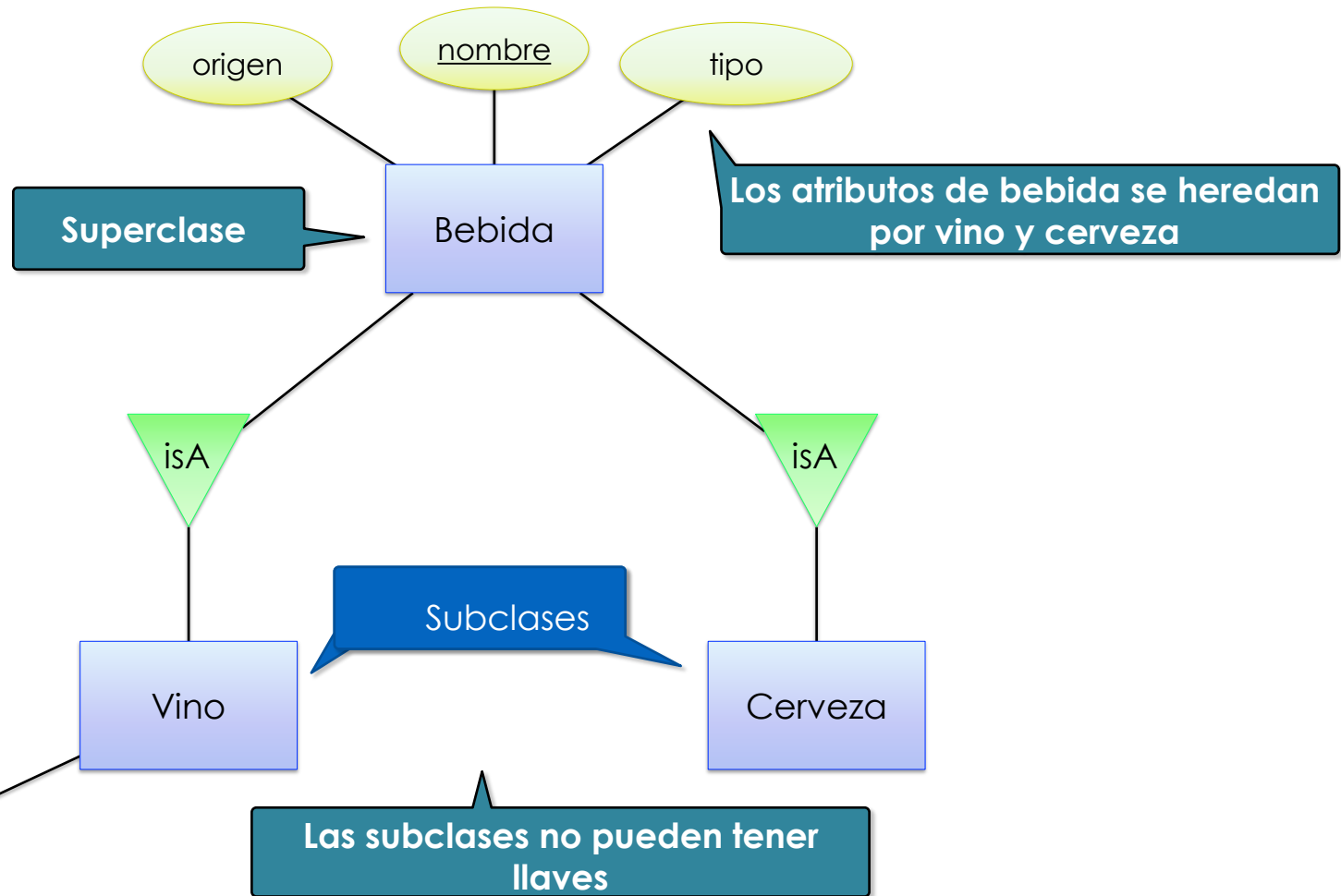


Diagramas E/R

Jerarquía de clases

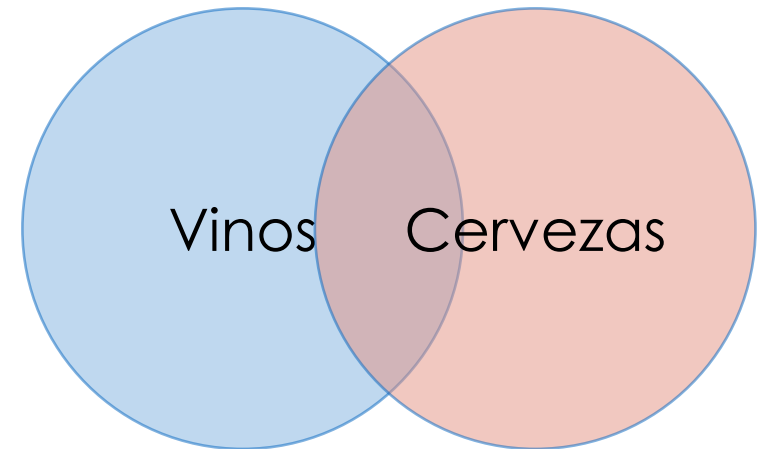
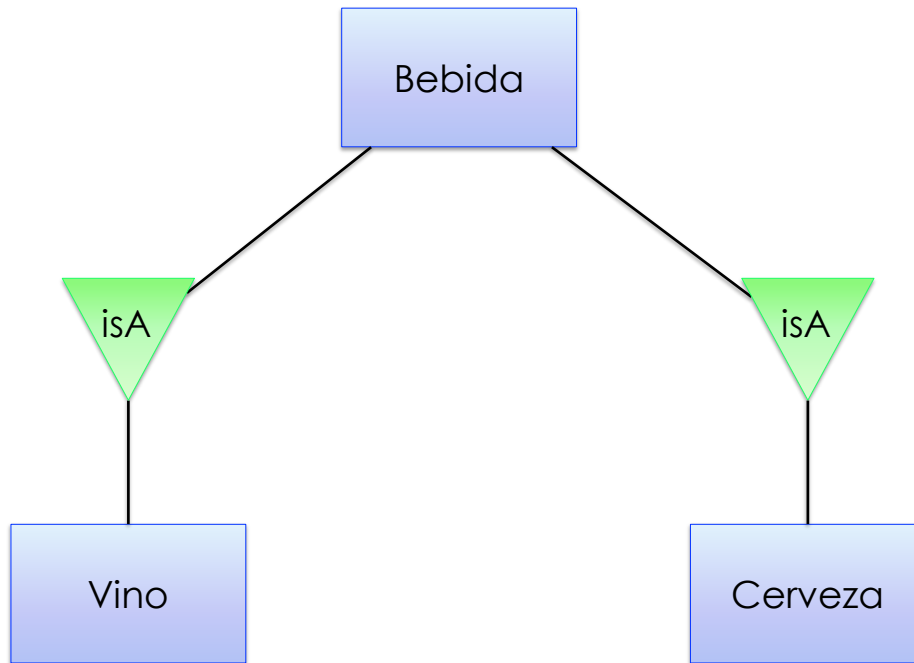


Jerarquía de clases



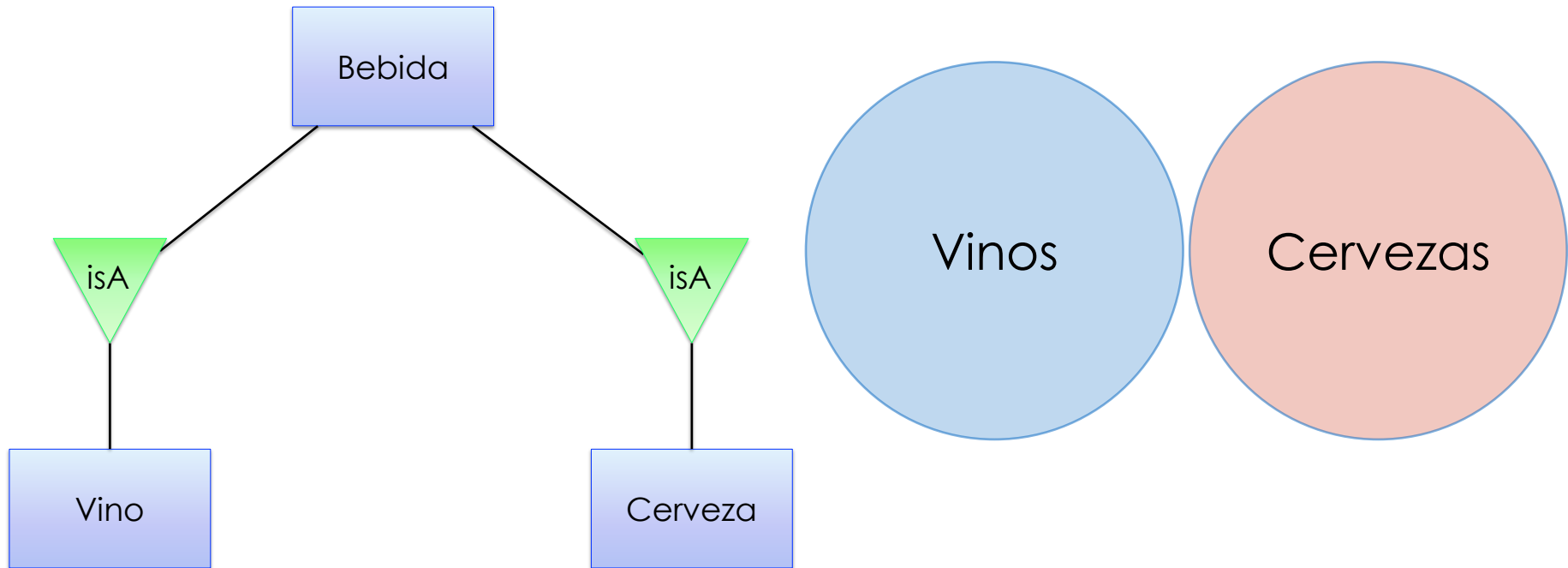
Jerarquía de clases

Solapamiento



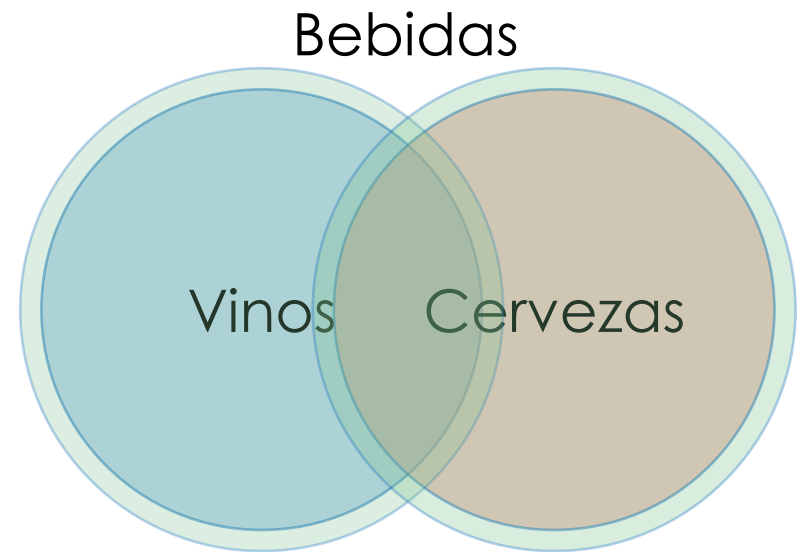
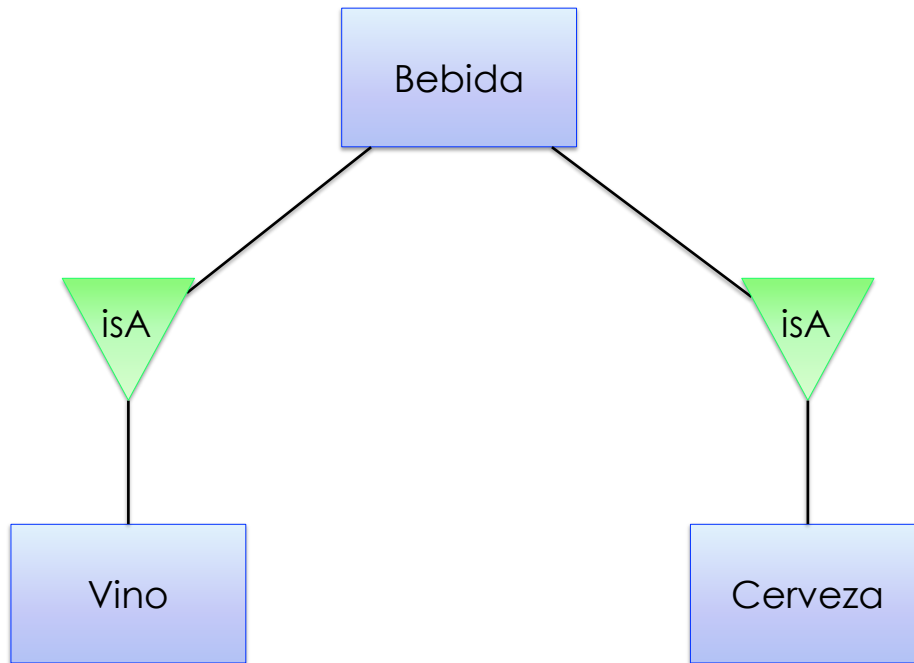
Jerarquía de clases

No Solapamiento



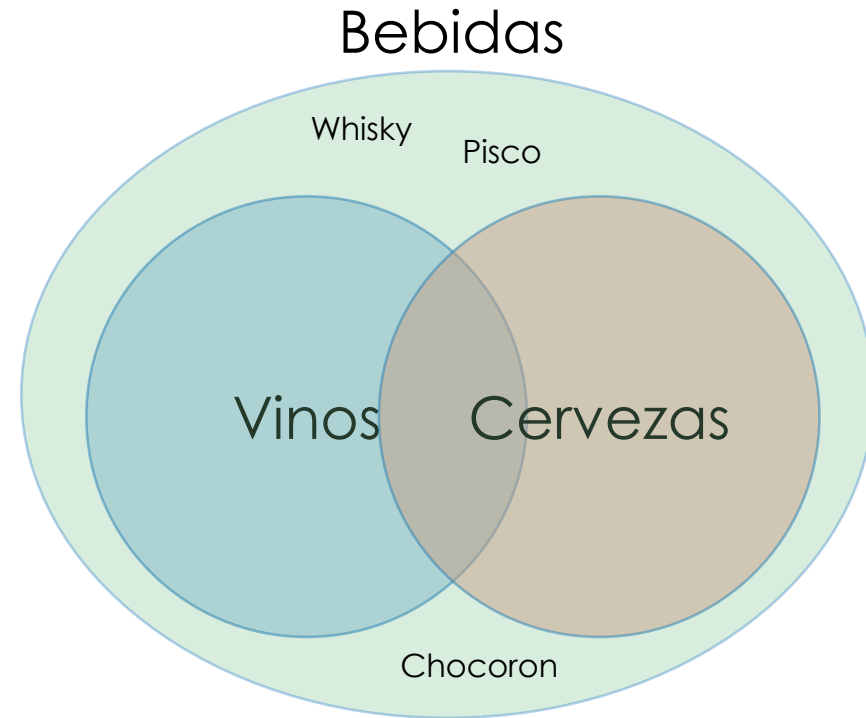
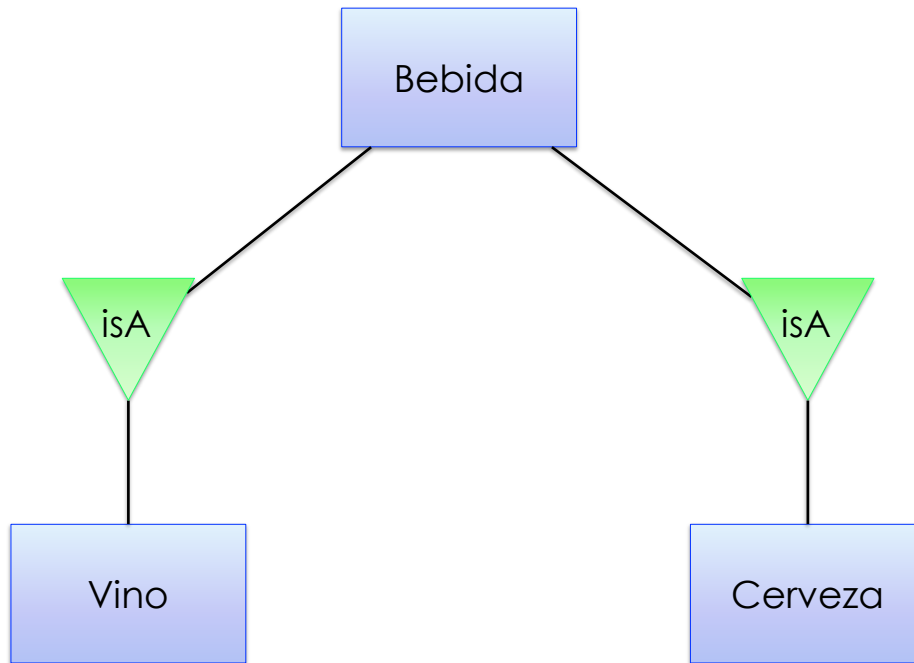
Jerarquía de clases

Cobertura

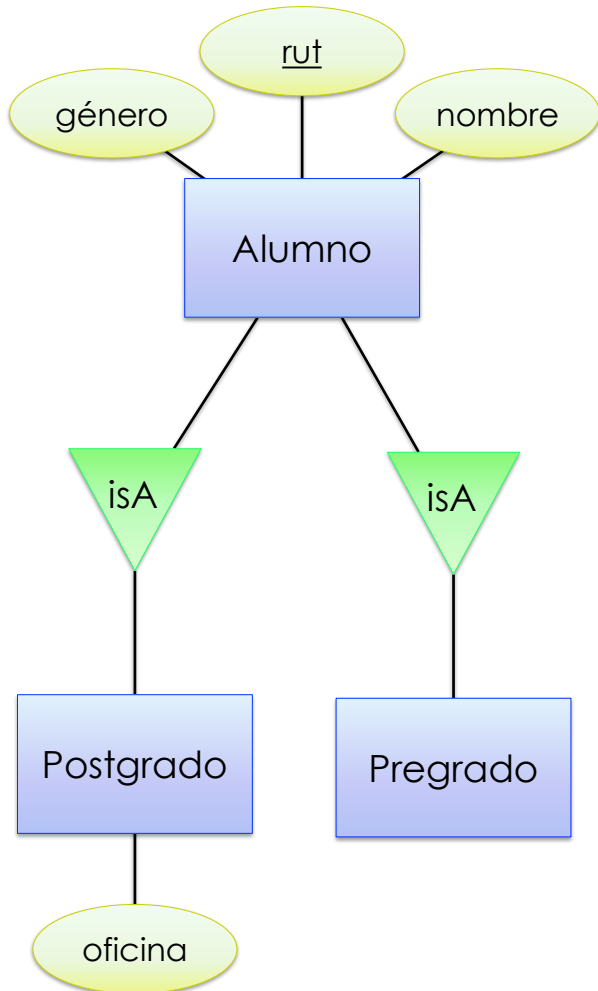


Jerarquía de clases

Sin Cobertura



Ejemplo



¿Hay **solapamiento**?

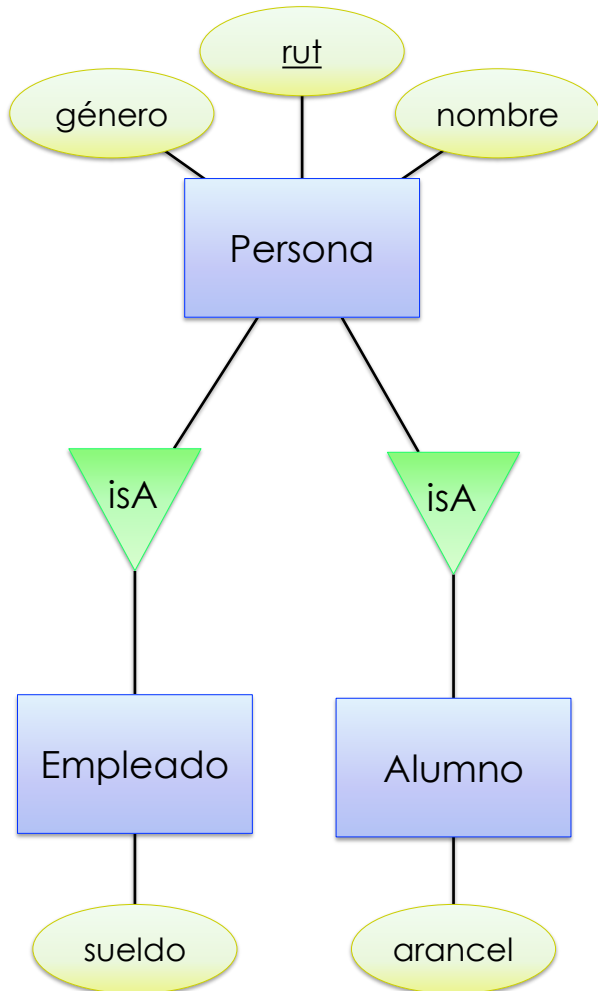
Depende (¿datos históricos?)

¿Hay **cobertura**?

Sí (de alumnos universitarios)



Ejemplo



¿Hay **solapamiento**?

Sí (profesor ayudante)

¿Hay **cobertura**?

Depende (¿visitantes?)

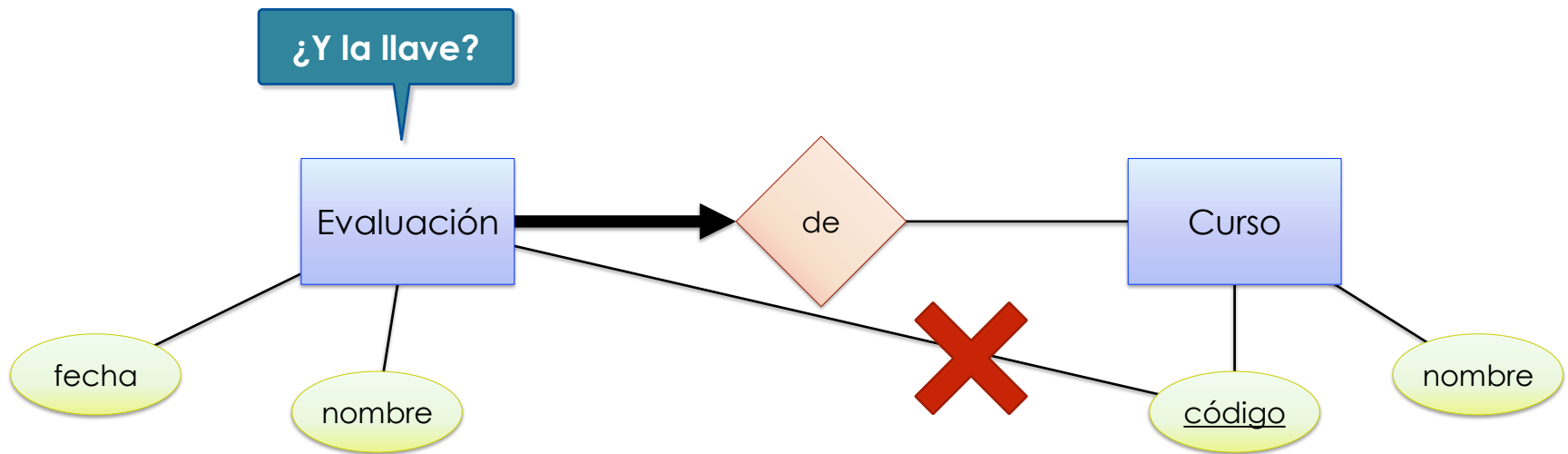


Diagramas E/R

Entidades Débiles

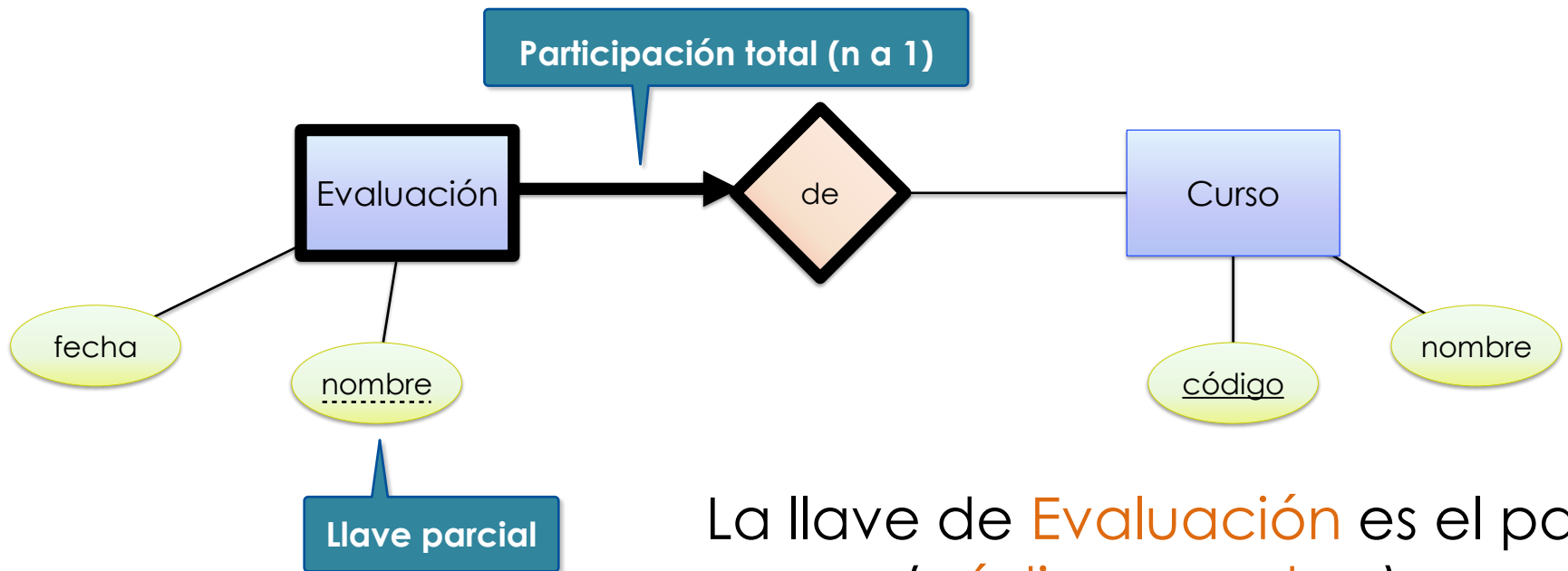


Entidades Débiles



Entidades Débiles

...entidades cuya llave dependa de la llave de otra entidad



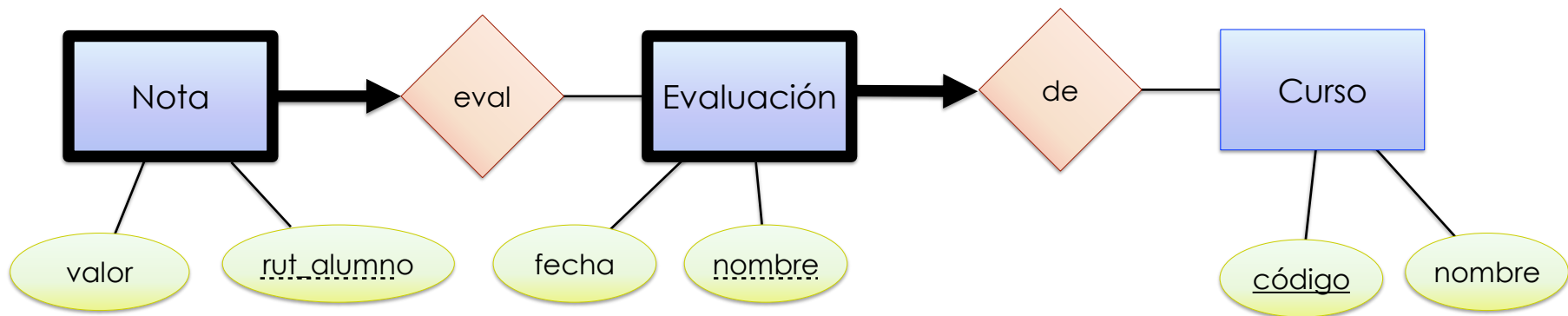
La llave de **Evaluación** es el par
(**código**, **nombre**)

¿Y las notas?...



Entidades Débiles

Podemos encadenar entidades débiles

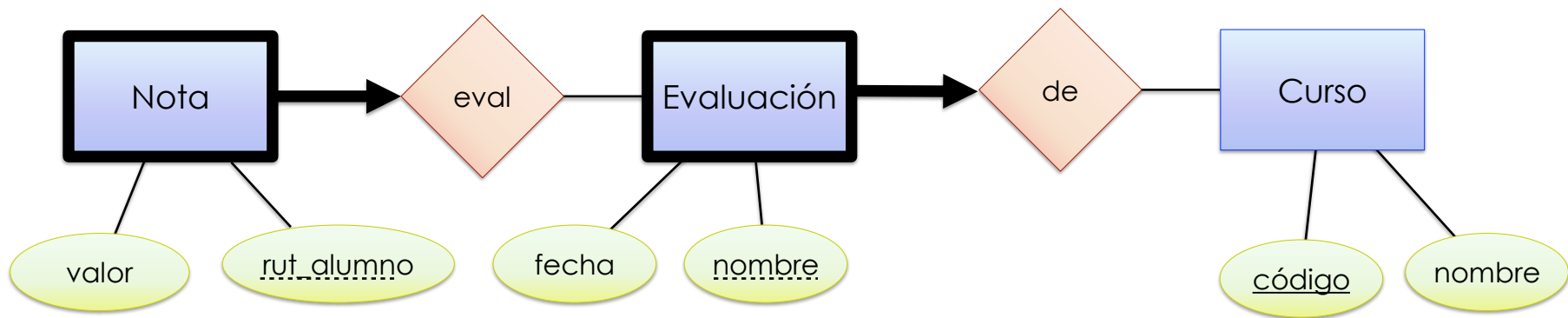


La llave de **Nota** es la tupla
(**código**, **nombre**, **rut_alumno**)



Entidades Débiles

Podemos encadenar entidades débiles



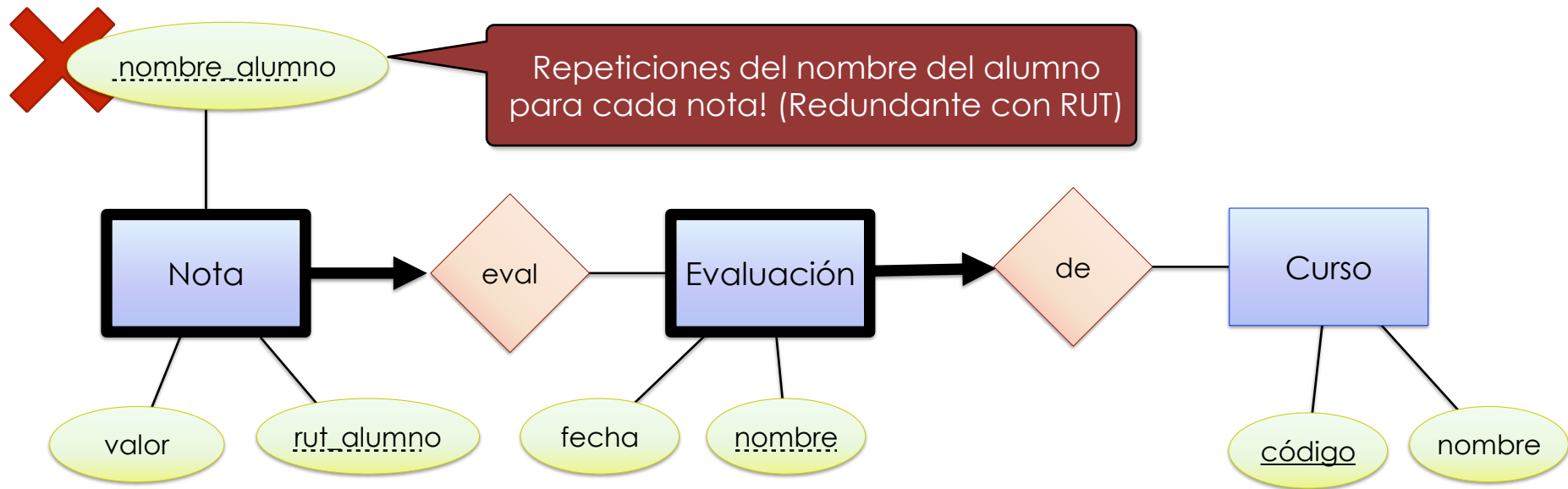
La llave de **Nota** es la tupla
(**código**, **nombre**, **rut_alumno**)

¿Y el alumno?...



Entidades Débiles

Agregando nombre del alumno

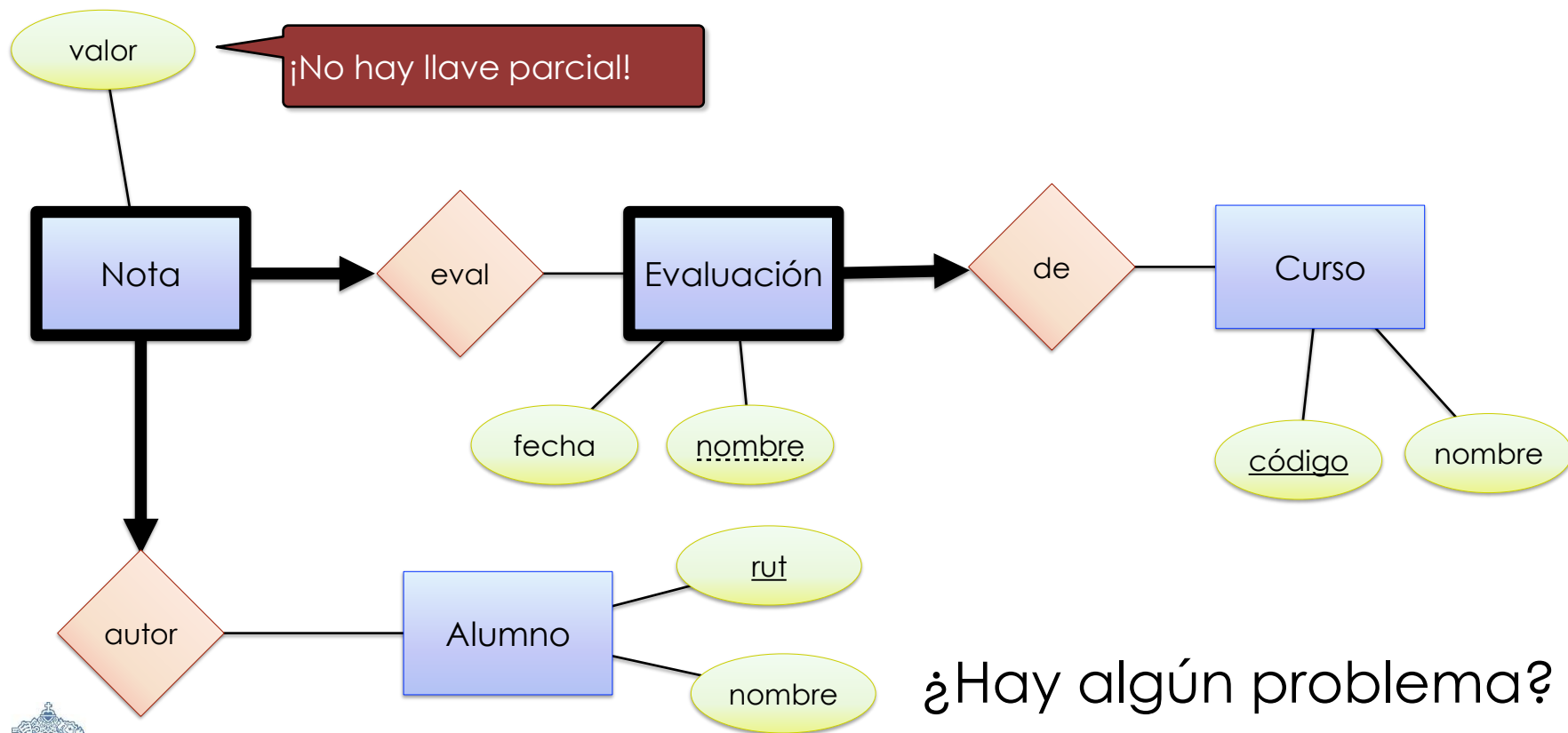


¿Hay algún problema?



Entidades Débiles

Agregando nombre del alumno

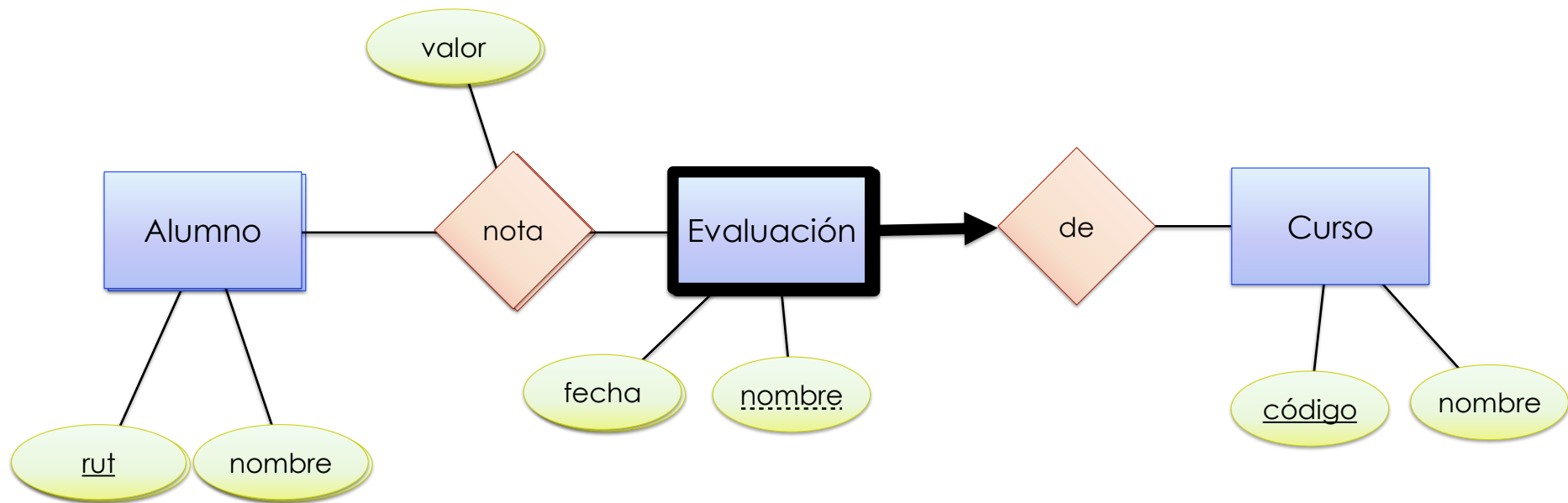


¿Hay algún problema?



Entidades Débiles

Agregando nombre del alumno



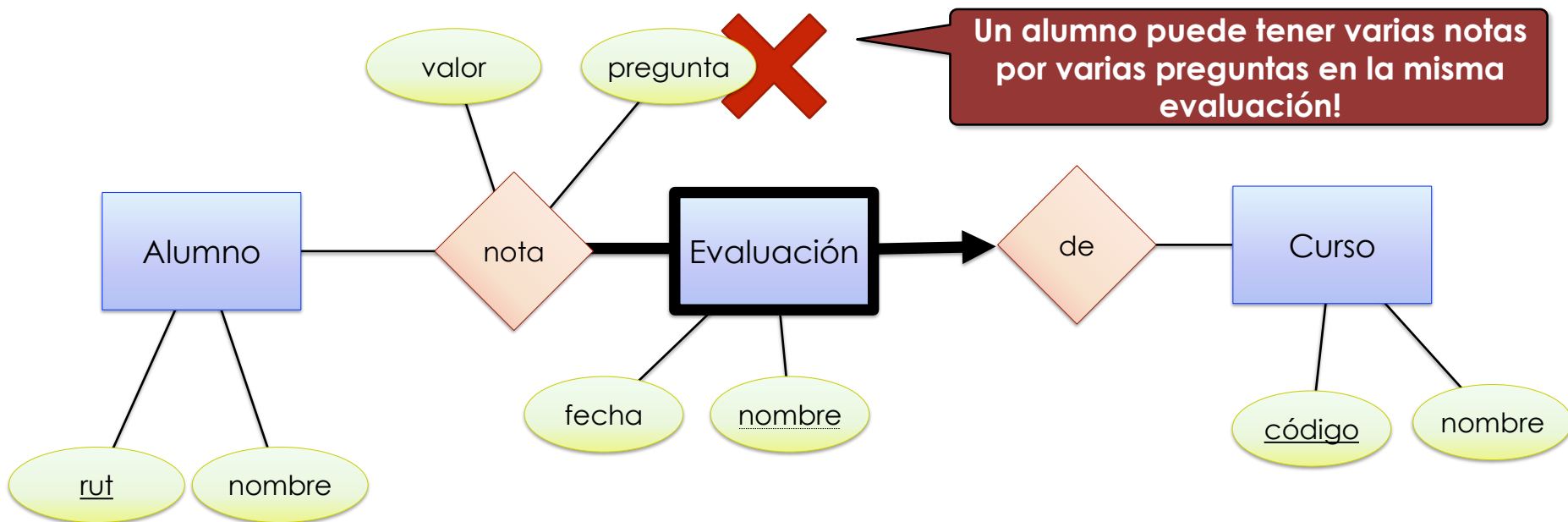
¿Hay algún problema?

¿Y notas por pregunta?...



Entidades Débiles

Nota por pregunta

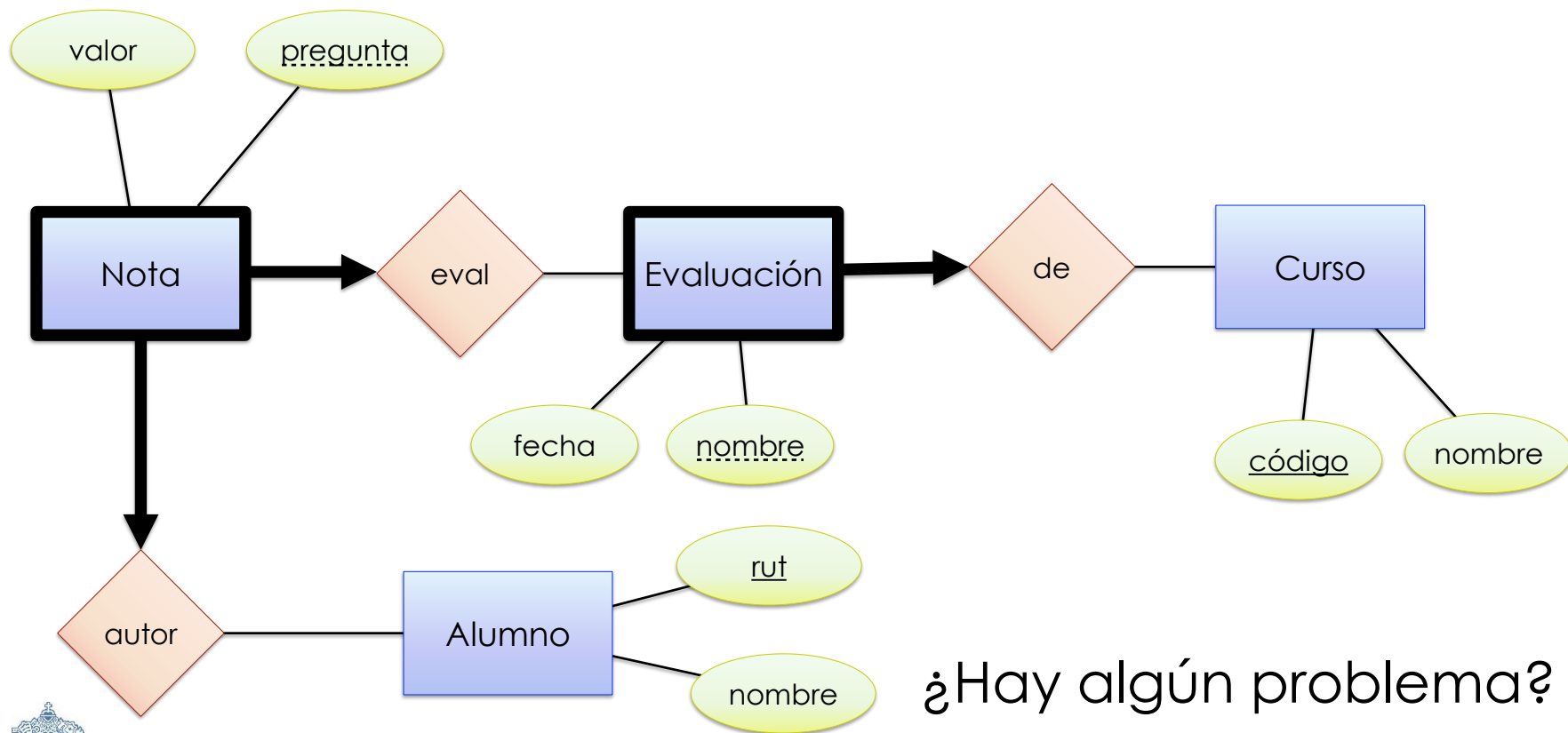


¿Hay algún problema?



Entidades Débiles

Nota por pregunta



¿Hay algún problema?

No :)



Diagramas E/R

Modelando vinos y cervezas



Modelando vinos y cervezas

Vendemos **vinos** y **cervezas**. Cada vino tiene año, tipo, grados y ciudad-origen. Cada cerveza tiene ciudad-origen, tipo, grados. Vinos y cervezas tienen un precio unitario y una cantidad “**en stock**” cada día.

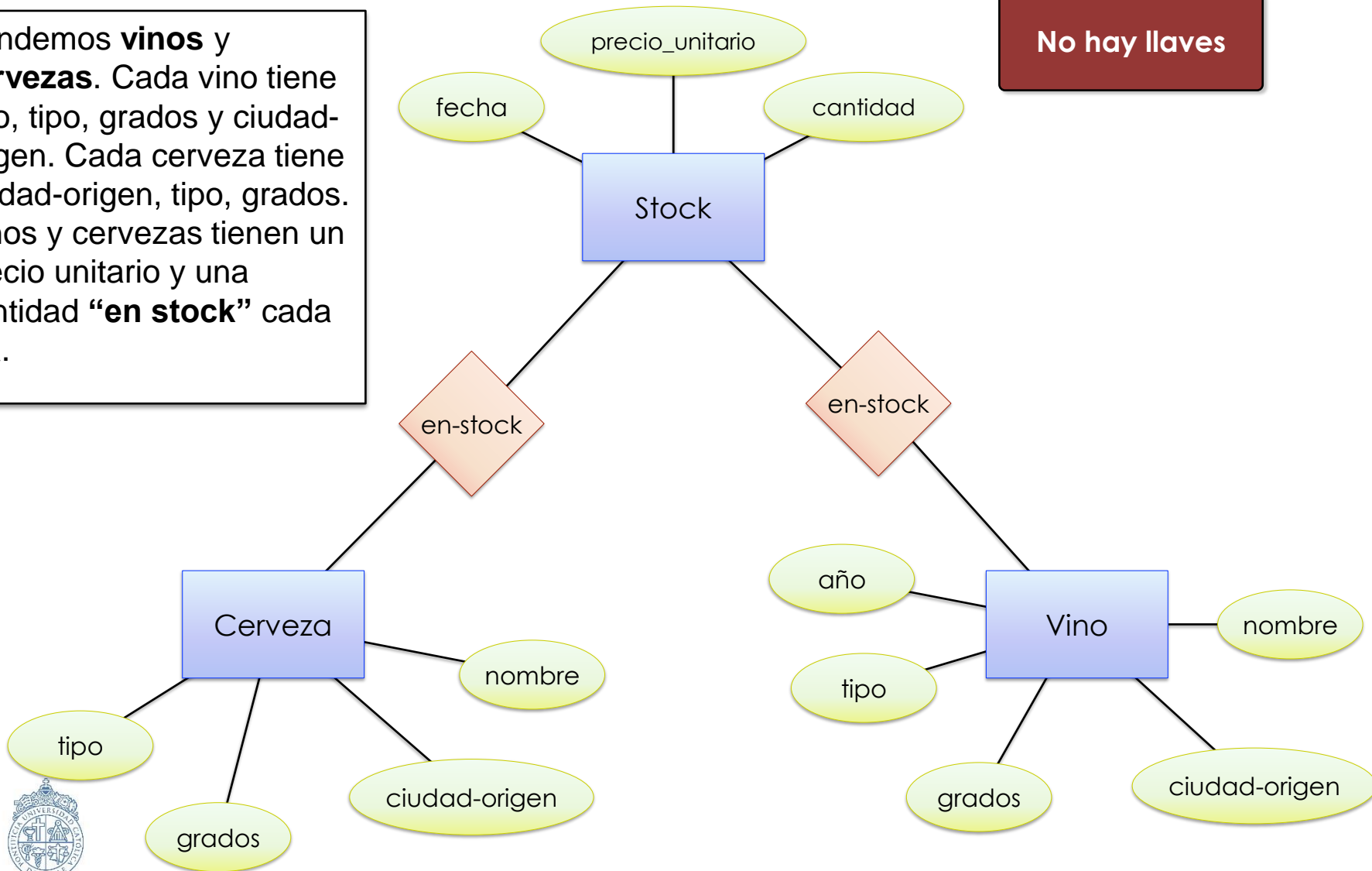


Modelando vinos y cervezas

Vendemos **vinos** y **cervezas**. Cada vino tiene año, tipo, grados y ciudad-origen. Cada cerveza tiene ciudad-origen, tipo, grados. Vinos y cervezas tienen un precio unitario y una cantidad “**en stock**” cada día.

¿Hay algún problema?

No hay llaves

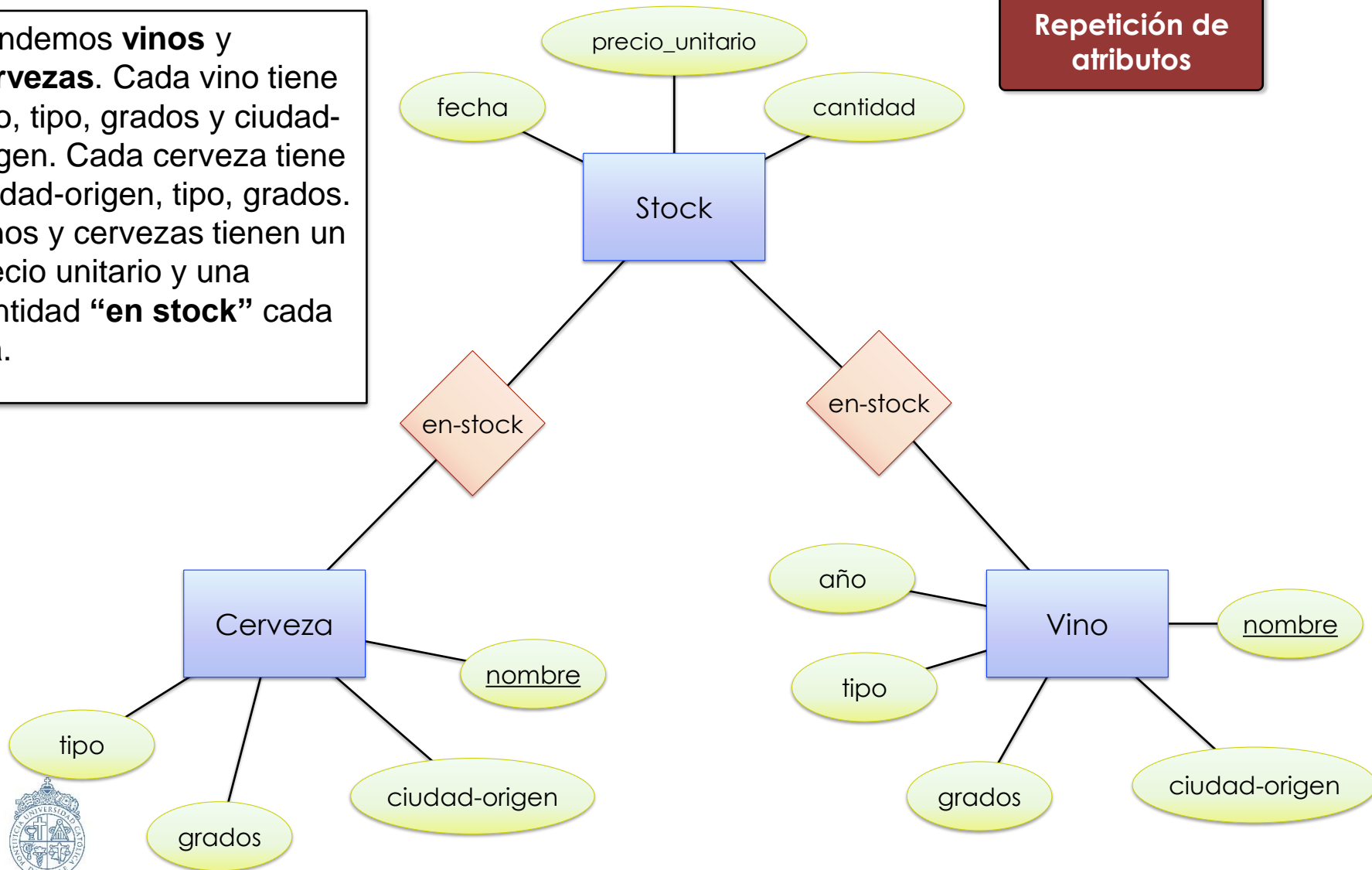


Agregando llaves

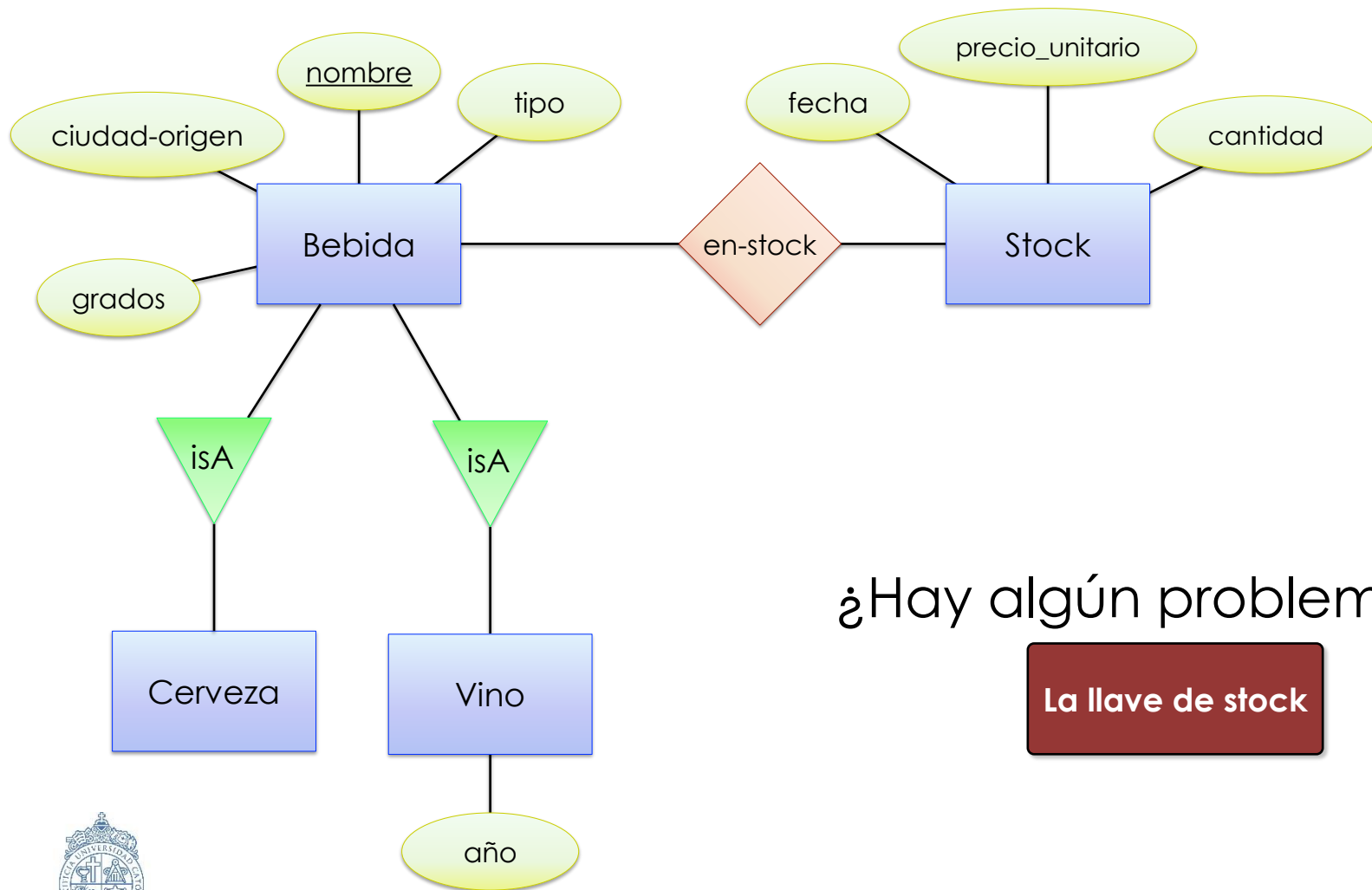
Vendemos **vinos** y **cervezas**. Cada vino tiene año, tipo, grados y ciudad-origen. Cada cerveza tiene ciudad-origen, tipo, grados. Vinos y cervezas tienen un precio unitario y una cantidad “**en stock**” cada día.

¿Hay algún problema?

Repetición de atributos



Usando jerarquía de clases

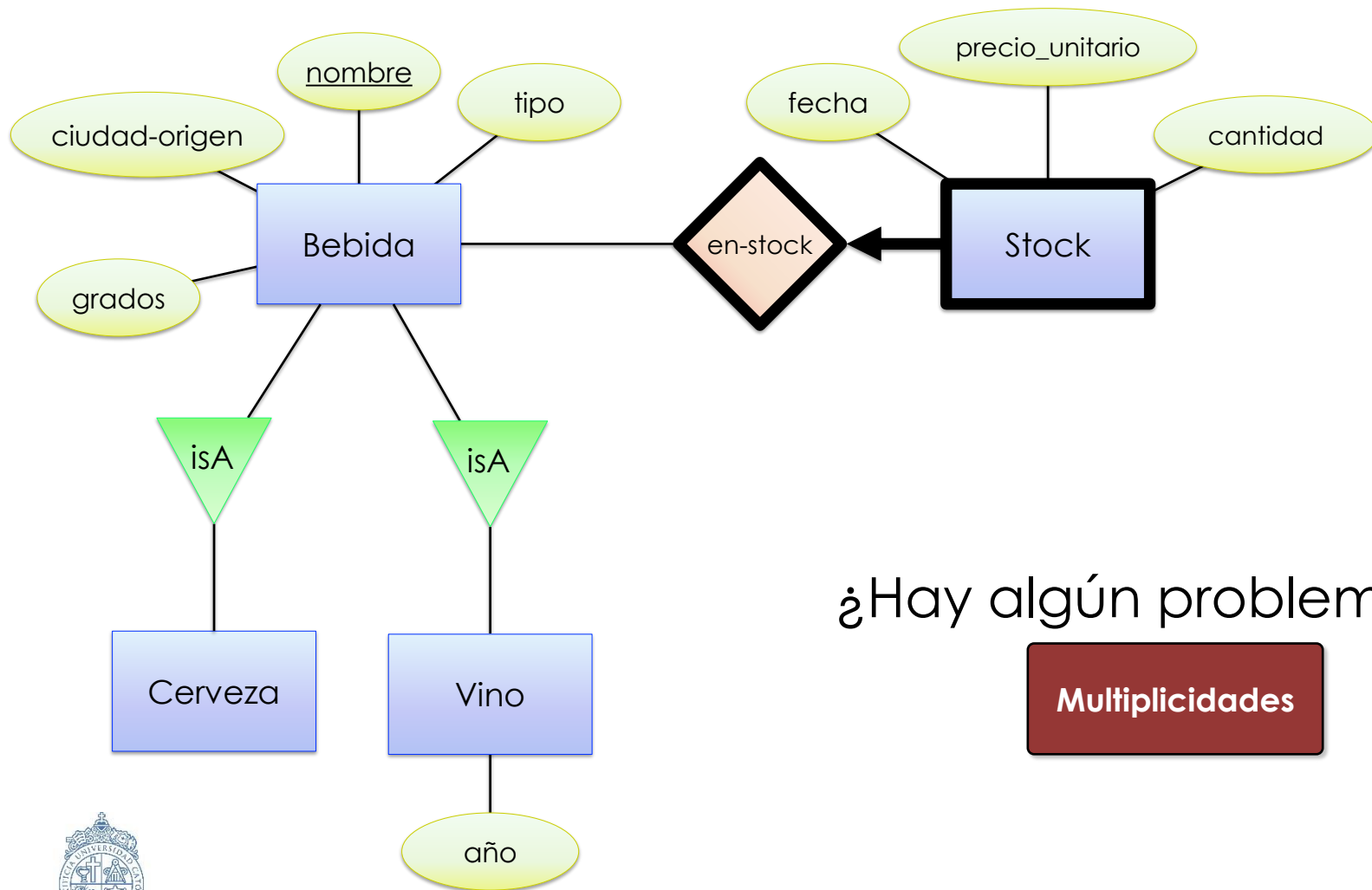


¿Hay algún problema?

La llave de stock



Usando jerarquía de clases

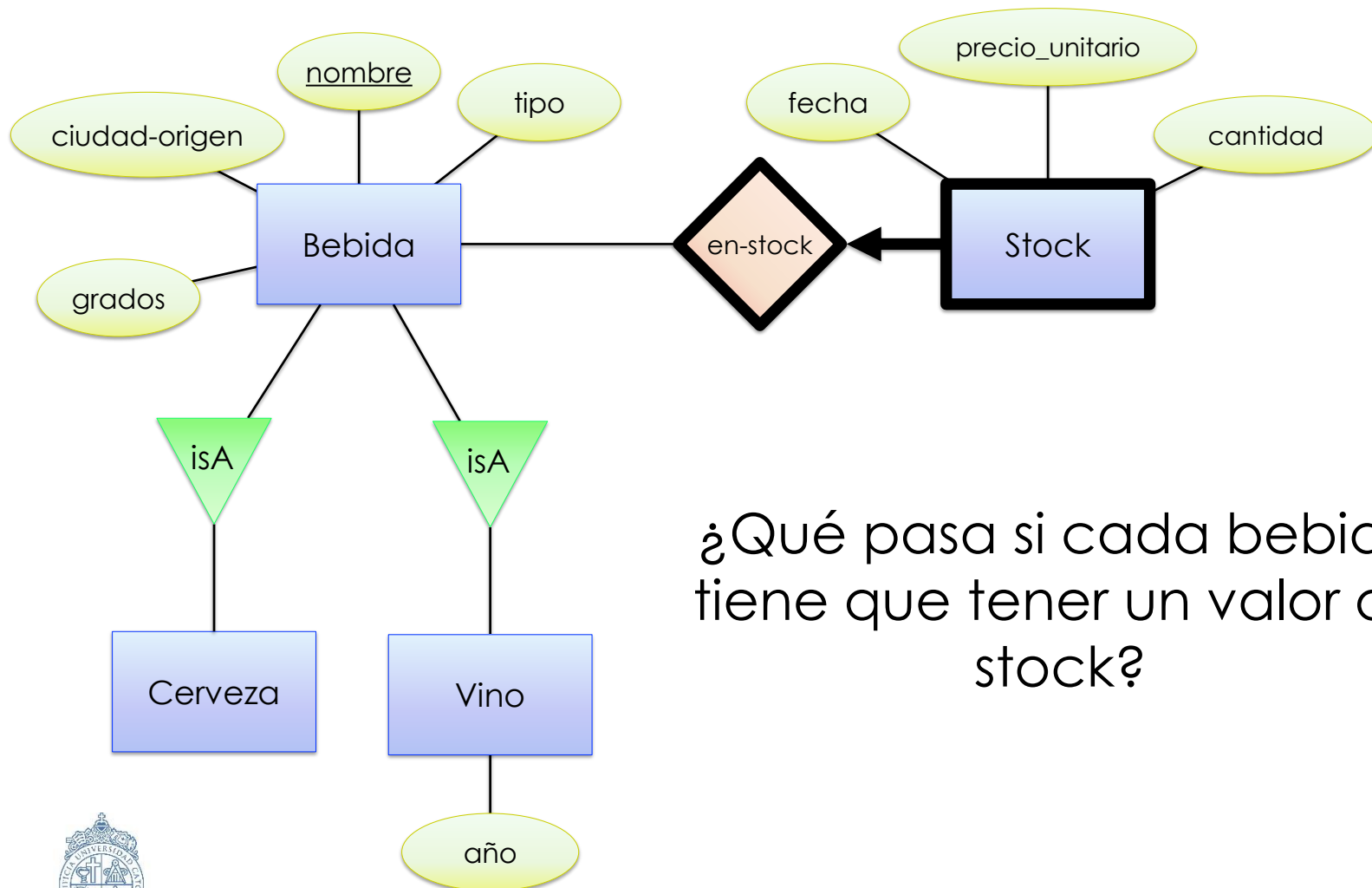


¿Hay algún problema?

Multiplicidades



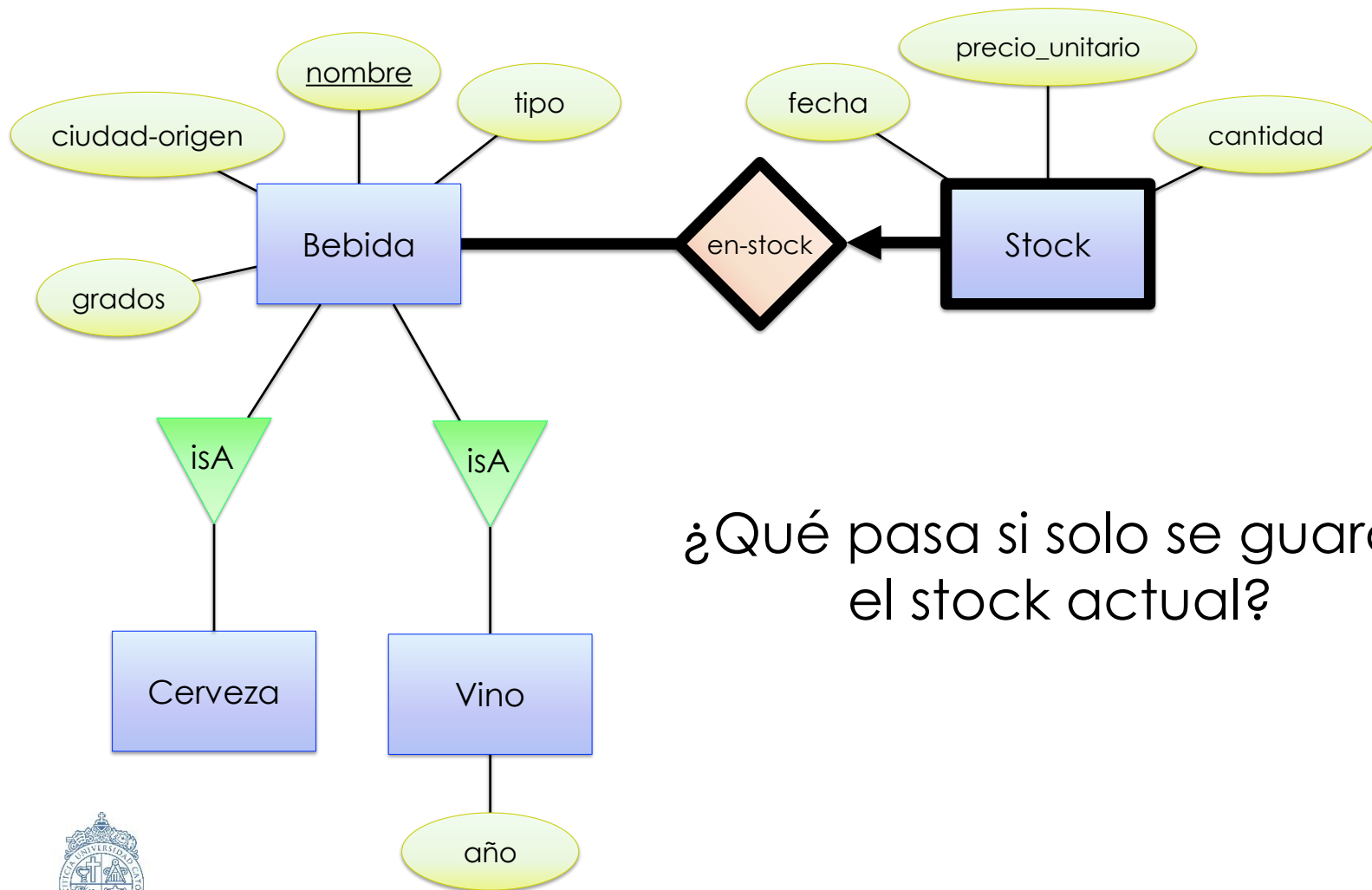
Agregando multiplicidades



¿Qué pasa si cada bebida tiene que tener un valor de stock?



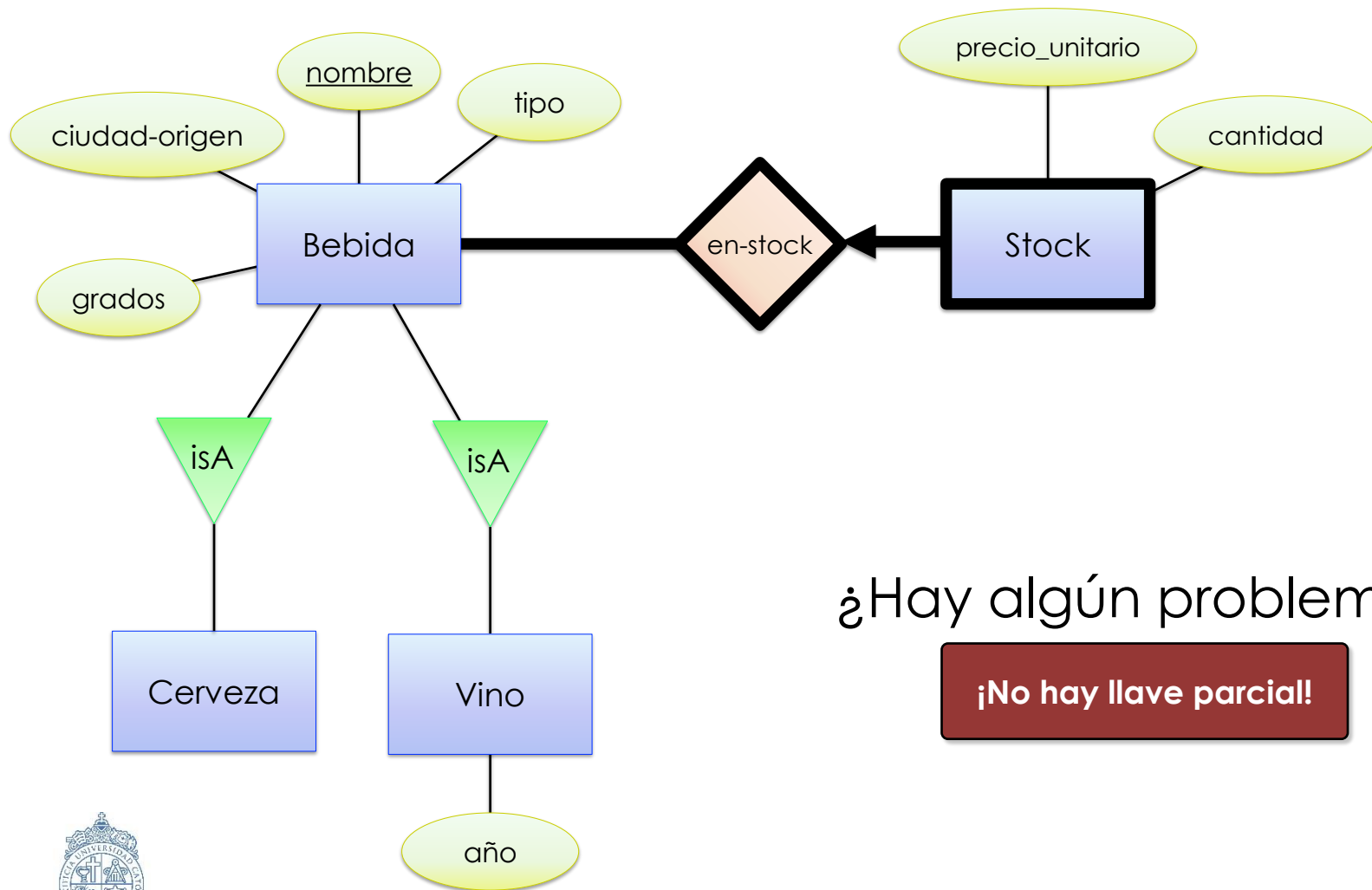
Agregando multiplicidades



¿Qué pasa si solo se guarda el stock actual?



Con sólo stock actual

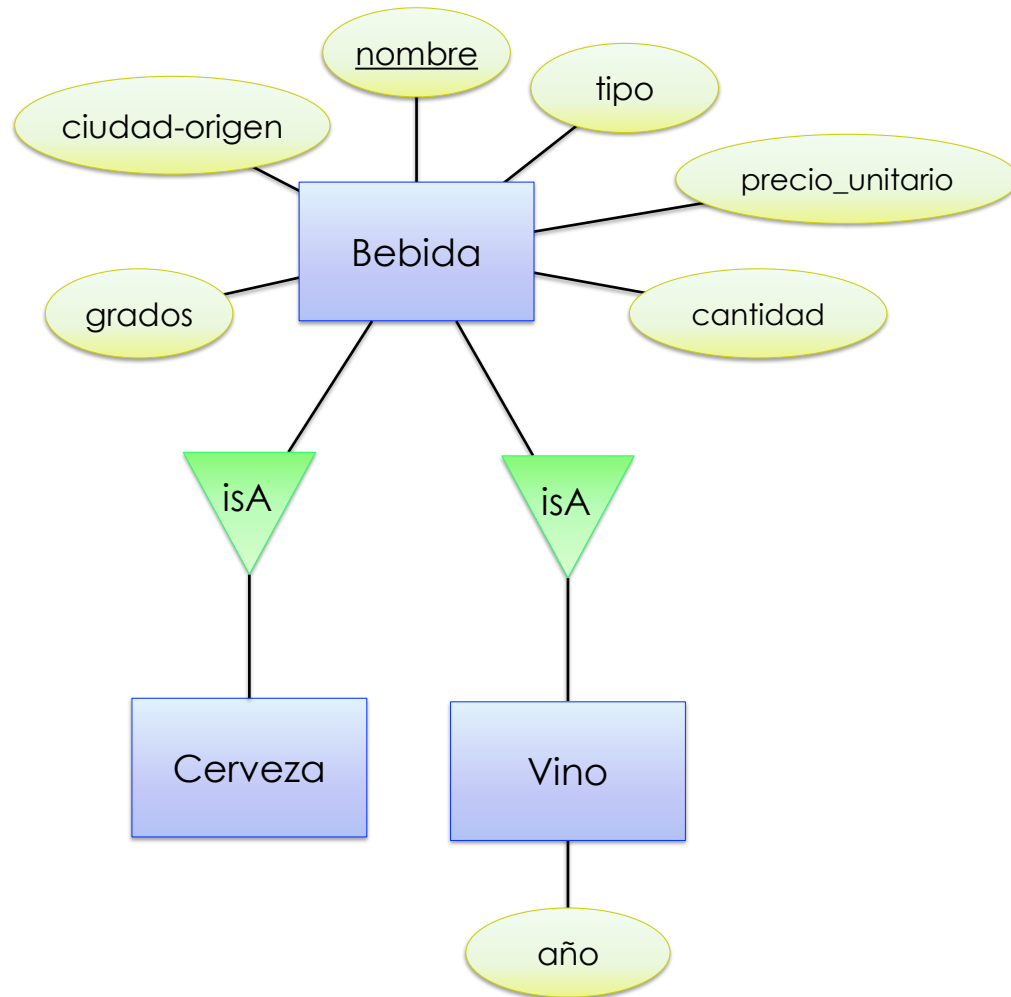


¿Hay algún problema?

¡No hay llave parcial!



Con sólo stock actual



¿Preguntas?

