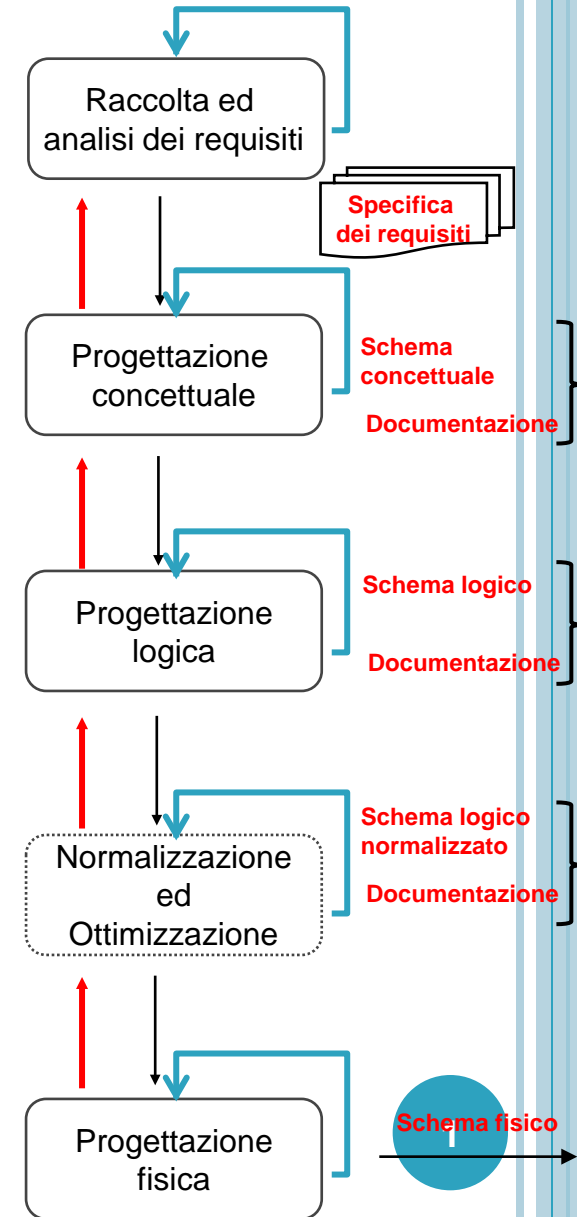




PROGETTAZIONE DI BASI DI DATI

PROGETTAZIONE

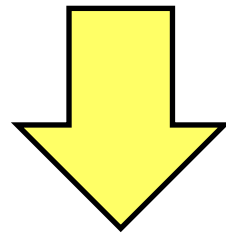
- Progettare una base di dati =
 - quali entità e associazioni rappresentare
 - come rappresentarle
 - Basi di dati
 - sempre più complesse
 - di dimensioni sempre più elevate
- ⇒ indispensabile approccio sistematico alla progettazione
- Attività di progettazione strutturata in fasi
 - modellazione della realtà per approssimazioni successive
 - Modello = rappresentazione formale della base di dati
 - astrae da aspetti non rilevanti per l'analisi che si vuole fare
 - permette di verificare specifiche
 - facilita comunicazione con gli utenti
 - Ogni fase prevede verifiche di qualità
 - Se la qualità non è sufficiente
 - **perfezionamenti o totali riscritture del modello generato in quella fase**
 - **ripetizione parziale o totale della fase precedente**
 - nessun effetto diretto sulle fasi non immediatamente collegate



RACCOLTA ED ANALISI DEI REQUISITI



documento in
linguaggio
naturale



caratteristiche
della base di dati

specifica dei requisiti

TIPOLOGIE DI REQUISITI

- **Requisiti informativi**
caratteristiche e tipologie dei dati
- **Requisiti sulle operazioni**
esplicitati nel carico di lavoro
- **Requisiti sui vincoli di integrità ed autorizzazione**
proprietà da assicurare ai dati, in termini di correttezza e protezione
- **Requisiti sulla popolosità della base di dati**
volume dei dati

PROGETTAZIONE CONCETTUALE

Dal documento di specifica si crea uno

schema concettuale

- descrizione formale ed ad alto livello
- del tutto indipendente dall'implementazione della base di dati
 - indipendente anche dal tipo di DBMS che sarà utilizzato - relazionale, object-relational, ecc.
- definito tramite un opportuno **modello concettuale**

PROGETTAZIONE CONCETTUALE

L'output di questa fase è

- Lo schema concettuale sviluppato utilizzando il modello concettuale prescelto
- Documentazione a corredo dello schema
 - Vincoli d'integrità non rappresentabili nello schema
 - Scelte progettuali
 - Altre informazioni che vedremo in seguito

PROGETTAZIONE LOGICA

- Traduzione dello schema concettuale nel modello dei dati del DBMS target
- L'output è lo **schema logico** nel DDL del DBMS target
- Aspetti considerati durante la progettazione logica:
 - Integrità e consistenza
 - Sicurezza
 - Efficienza

NORMALIZZAZIONE

- Verifica di qualità dello schema logico prodotto, effettuata tramite opportuni strumenti formali
- Prende il nome di normalizzazione nel caso di basi di dati relazionali

PROGETTAZIONE FISICA

- In questa fase vengono effettuate alcune scelte circa la memorizzazione fisica dei dati (ad esempio, indici)
- L'output è lo **schema fisico** che descrive le strutture di memorizzazione ed accesso ai dati

METODOLOGIE DI PROGETTAZIONE

- Il modello ER è lo strumento tramite cui è possibile definire lo schema concettuale di una base di dati, ma...
- Come ottenere un “buon” schema concettuale partendo dal documento di specifica dei requisiti?
- Come si valuta la “bontà” di uno schema concettuale?
- Non esiste una metodologia totalmente standardizzabile ... ma esistono alcune linee guida per condurre le varie fasi

RACCOLTA ED ANALISI DEI REQUISITI

- È una attività difficilmente standardizzabile
- Di solito condotta da personale esperto, in stretta collaborazione con gli utenti finali
- Una volta ottenuto il documento di specifica, è però possibile effettuare alcune operazioni su di esso, che faciliteranno le successive fasi di progettazione

RACCOLTA ED ANALISI DEI REQUISITI

OPERAZIONI SUL DOCUMENTO DI SPECIFICA

- Leggerlo attentamente evidenziando
 - tutte le possibili fonti di ambiguità ed imprecisione
 - sinonimi usati per lo stesso concetto
- Chiarire i dubbi con i committenti del progetto
- Scegliere per ogni concetto il termine che meglio lo modella e usare esclusivamente quello
- Separare le frasi che riguardano i dati da quelle che riguardano le operazioni
- Riorganizzare le frasi per concetti

PROGETTAZIONE CONCETTUALE

- Ristrutturato il documento di specifica, come è possibile ottenere da questo un buon schema ER?
 - Scelta dei costrutti
 - Metodologie per la generazione di diagrammi ER
 - Documentazione dei diagrammi
 - Verifiche di qualità

SCelta DEI CoSTRUTTI

- Leggere attentamente il documento di specifica e per ogni concetto individuato, scegliere il costrutto del modello ER più idoneo alla sua rappresentazione
- Esistono alcune linee guida per effettuare questa scelta

SCelta DEI CoSTRUTTI: ENTITÀ

- Per rappresentare concetti che descrivono oggetti
 - omogenei, cioè caratterizzati da un insieme di proprietà comuni
 - rilevanti per il dominio considerato
- Es. clienti della videoteca, film
- Candidati tipici i **nomi** che compaiono frequentemente nel documento di specifica
- Dal documento di specifica si deducono anche i vincoli di identificazione delle entità

SCelta DEI CoSTRUTTI:

GERARCHIA DI GENERALIZZAZIONE

- Per rappresentare concetti che sono un caso particolare di un altro
- Es. clienti, clienti standard e clienti VIP
- Si usa quando le entità figlie hanno proprietà in più rispetto all'entità padre
 - es. Clienti standard e VIP hanno ciascuno un attributo (diverso) per la gestione dei punti
- Altrimenti si inserisce nell'entità che modella il concetto più generale un attributo che indentifica le varie tipologie di istanze
 - es. video con attributo tipo per indicare se DVD o VHS

SCelta DEI CoSTRUTTI: ATTRIBUTO

- Si usa per rappresentare una proprietà di un concetto modellato come entità
- Es. nome di un cliente, titolo di un film, codice di un video
- Per modellare proprietà con sotto-proprietà
 - Es. residenza dei clienti della videoteca <Via, numero civico, città, cap>
 - attributi composti \Rightarrow duplicazione di eventuali dati uguali in entità diverse (più clienti con stesso indirizzo)
 - oppure entità \Rightarrow eventuali dati uguali in entità diverse in copia unica
 - la seconda si usa quando lo stesso valore dell'attributo si ripete in più entità
- Dal documento di specifica si deducono anche i vincoli di cardinalità

SCelta DEI CoSTRUTTI: ASSOCIAZIONE

- Si usa per rappresentare un legame logico tra concetti modellati come entità
- Es. video consigliato da un cliente
- Una associazione può anche avere delle proprietà, modellate come attributi
- Candidati tipici i verbi che mettono in relazione concetti modellati come entità
- Dal documento di specifica si deducono anche i vincoli di cardinalità

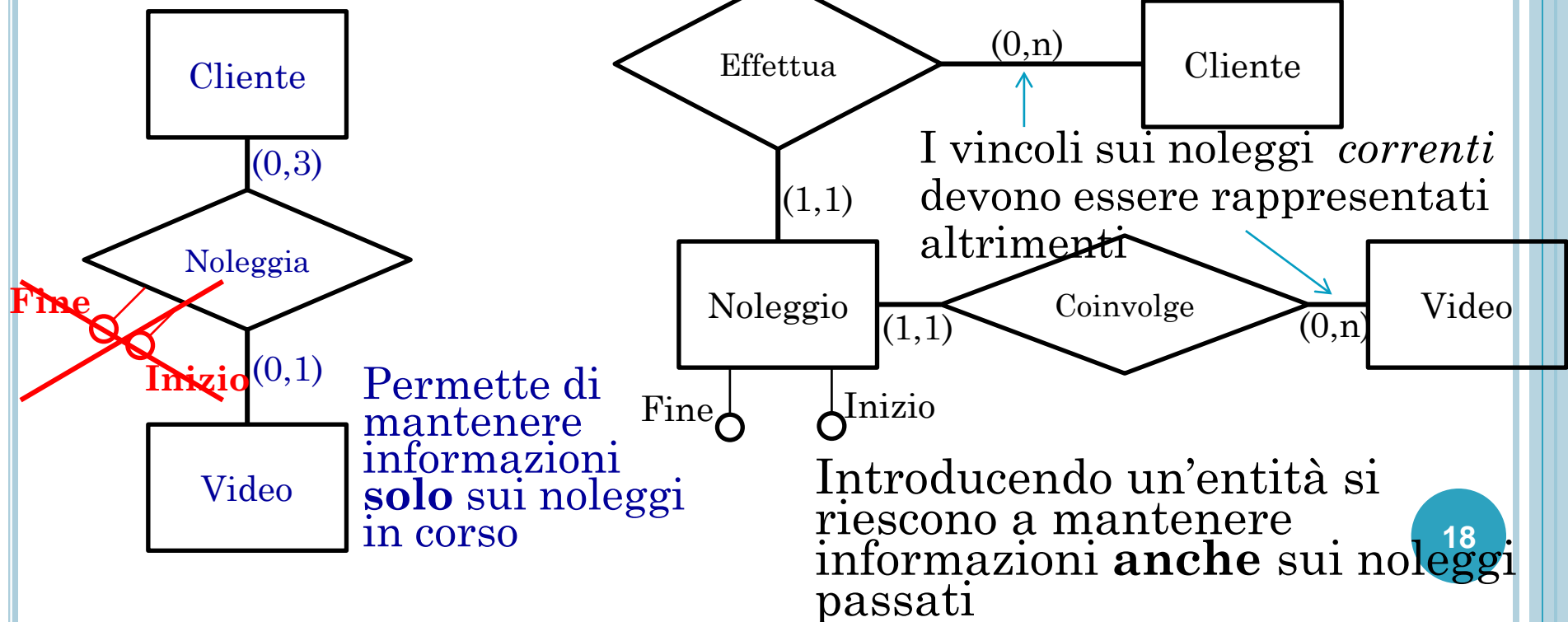
SCELTA DEI COSTRUTTI

NON SEMPRE COSÌ OVVIA

- Non sempre è facile decidere se modellare un concetto tramite un'entità o una associazione

Esempio

- Un cliente noleggia fino a 3 video contemporaneamente



ALTERNATIVE

- Esercizio proposto di modellazione alternativa
 - usare un'associazione ternaria introducendo l'entità DatiNoleggio
 - usare l'associazione di sinistra per i noleggi correnti ed un'altra per mantenere lo storico
 - di quali operazioni in più ho bisogno per gestire questa soluzione rispetto alle altre?
 - che cosa si semplifica a livello di vincoli

GENERAZIONE DIAGRAMMI ER

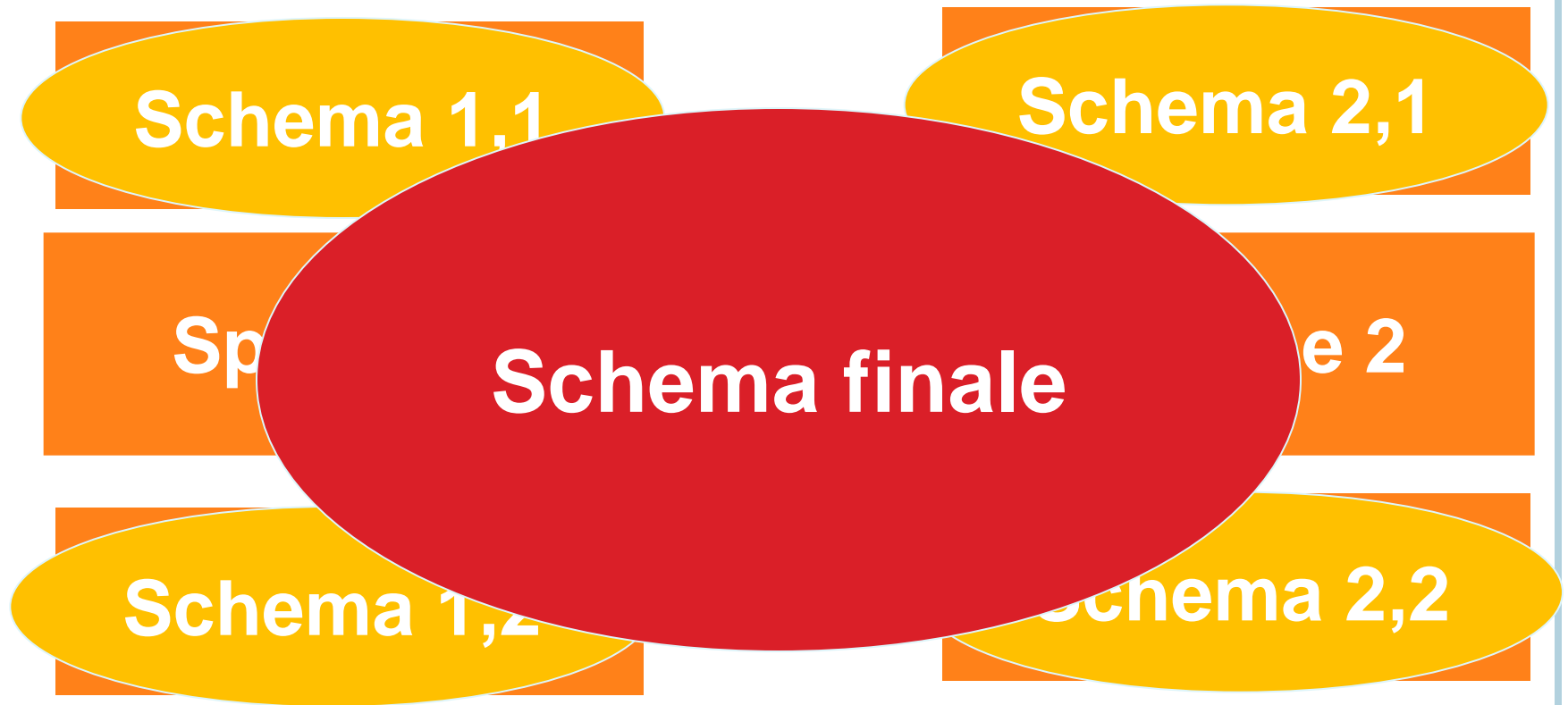
- Tecniche base per generare un diagramma ER *finale*
 - raffinamento
 - integrazione di sottoschemi
- Strategie
 - Top-down (prevale raffinamento)
 - Bottom-up (prevale suddivisione in sottoschemi)
 - Mista

STRATEGIA TOP-DOWN



Schema finale

STRATEGIA BOTTOM-UP



STRATEGIA MISTA

- Si individuano i concetti principali e si realizza uno **schema scheletro**
- Sulla base di questo si decompone in sotto-problemi
- Ciascun sotto-problema viene
 - raffinato
 - espanso
- Le soluzioni dei sotto-problemi vengono integrate per arrivare allo schema finale

DOCUMENTAZIONE DI SUPPORTO

Deve contenere

- Domini degli attributi
- Vincoli imposti dal dominio non rappresentabili tramite vincoli di identificazione e cardinalità nel diagramma ER
- Vincoli di autorizzazione, se necessari
- Tipologie di gerarchie di generalizzazione
- Informazioni sulle principali scelte progettuali, soprattutto ove siano possibili più alternative
- Dizionario delle entità e delle associazioni, per basi di dati di dimensioni elevate

VERIFICHE DI QUALITÀ

- Frequenti verifiche di completezza e correttezza sugli schemi intermedi
- Verifiche sullo schema ER finale:
 - Correttezza sintattica e semantica
 - Completezza
 - Ridondanza

UN ESEMPIO DI PROGETTAZIONE CONCETTUALE

Videoteca

SPECIFICHE

Vogliamo realizzare una base di dati per una videoteca. La videoteca consente il noleggio di circa 1000 film. Per ogni film, vogliamo memorizzare il titolo, il regista, l'anno di produzione, il genere e la valutazione della critica, se presente. Ogni film è disponibile per il noleggio in un certo numero di video. Ogni videocassetta o dvd disponibile nella videoteca (circa 3000) è identificato da un codice di collocazione. La base di dati dovrà inoltre memorizzare informazioni sui clienti della videoteca (circa 2000) e sui video che hanno noleggiato. Il numero di noleggi giornalieri alla videoteca è circa 200. Per ogni utente della videoteca vogliamo mantenere il suo nome, cognome, data di nascita, residenza e telefono. Ogni cliente è identificato da un codice che corrisponde al numero della tessera rilasciatagli per usufruire dei servizi della videoteca. Ogni cliente può avere contemporaneamente in noleggio un certo numero di video (non più di tre). Per ogni noleggio, vogliamo memorizzare la data in cui il noleggio è stato effettuato e, per i noleggi conclusi, la data di restituzione. Ogni cliente può inoltre consigliare dei film ad altri clienti, esprimendo per essi un giudizio. La videoteca prevede un programma di fidelizzazione dei clienti. Ogni noleggio consente l'accumulo di un certo numero di punti. Quando i punti accumulati superano una certa soglia, i clienti vengono qualificati come clienti VIP e hanno diritto ad un bonus. La base di dati dovrà memorizzare, per ogni cliente VIP, il valore corrente del suo bonus. Per i clienti standard, ovvero quelli che non hanno ancora accumulato punti sufficienti per accedere alla categoria VIP, vogliamo memorizzare il numero di punti mancanti ad accedere a tale categoria

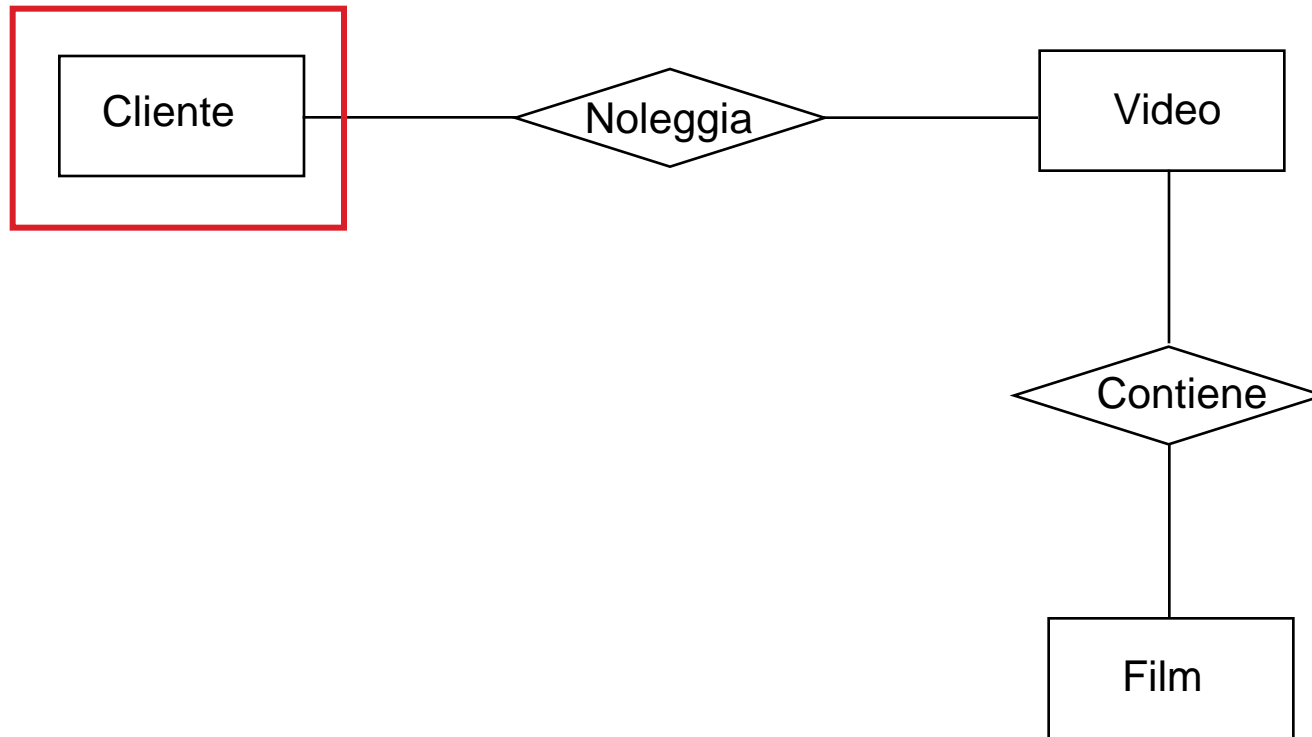
ANALISI DELLE SPECIFICHE

Vogliamo realizzare una base di dati per una videoteca. La videoteca consente il noleggio di circa 1000 film. Per ogni **film**, vogliamo memorizzare il titolo, **il regista**, l'anno di produzione, il genere e **la valutazione della critica**, se presente. Ogni film è disponibile per il noleggio in un certo numero di **video**. Ogni **videocassetta o dvd** disponibile nella videoteca (circa 3000) è identificato da un codice di collocazione. La base di dati dovrà inoltre memorizzare informazioni sui **clienti** della videoteca (circa 2000) e sui **video che hanno noleggiato**. Il numero di noleggi giornalieri alla videoteca è circa 200. Per ogni **utente** della videoteca vogliamo mantenere il suo nome, cognome, data di nascita, **residenza** e **telefono**. Ogni cliente è **identificato da un codice** che corrisponde al numero della tessera rilasciatagli per usufruire dei servizi della videoteca. Ogni cliente può avere contemporaneamente in noleggio un certo numero di video (non più di tre). Per ogni noleggio, vogliamo memorizzare **la data** in cui il noleggio è stato effettuato e, per i noleggi conclusi, **la data di restituzione**. Ogni cliente può inoltre consigliare dei film ad altri clienti, esprimendo per essi **un giudizio**. La videoteca prevede un programma di fidelizzazione dei clienti. Ogni noleggio consente l'accumulo di un certo numero di punti. Quando i punti accumulati superano una certa soglia, i clienti vengono qualificati come clienti VIP e hanno diritto ad un bonus. La base di dati dovrà memorizzare, per ogni cliente VIP, il valore corrente del **suo bonus**. Per i clienti standard, ovvero quelli che non hanno ancora accumulato punti sufficienti per accedere alla categoria vogliamo memorizzare il numero di punti mancanti ad accedere a tale categoria

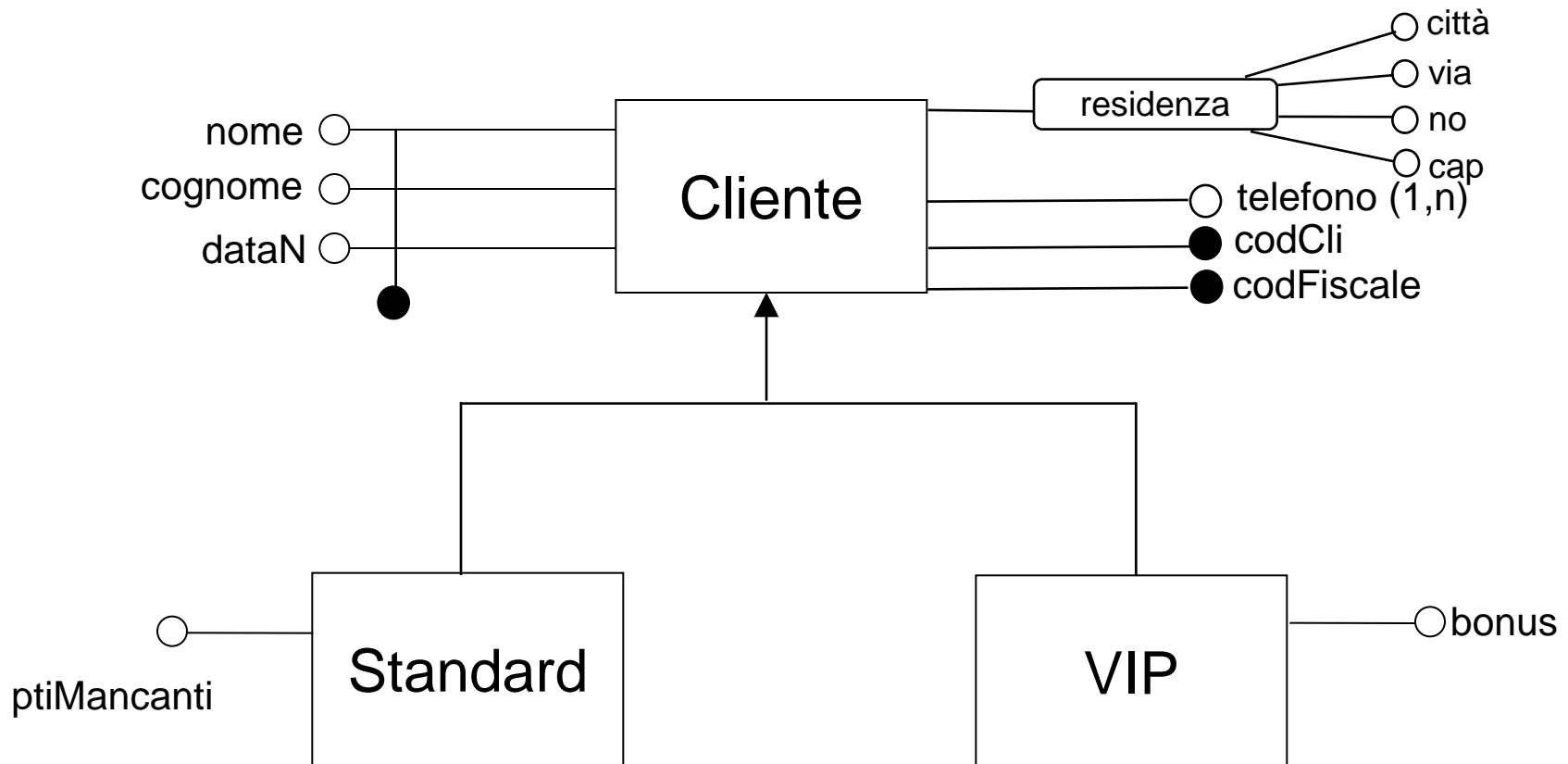
SPECIFICHE RISTRUTTURATE

Vogliamo realizzare una base di dati per una videoteca. La videoteca consente il noleggio di circa 1000 film. Per ogni film **(identificato da titolo e nome+cognome del regista)**, vogliamo memorizzare il titolo, **il nome e cognome del regista**, l'anno di produzione, il genere e la valutazione della critica **(espressa in una scala di valori decimali da 0 a 5)**, se presente. Ogni film è disponibile per il noleggio in un certo numero **di video (dove video sono sia videocassette che dvd)**. **Ogni video** disponibile nella videoteca (circa 3000) è identificato da un codice di collocazione. La base di dati dovrà inoltre memorizzare informazioni sui clienti della videoteca (circa 2000) e sui video che hanno **correntemente in noleggio ed hanno noleggiato in passato**. Il numero di noleggi giornalieri alla videoteca è circa 200. Per ogni **cliente** della videoteca vogliamo mantenere il suo nome, cognome, data di nascita, residenza **(intesa come città, via, numero civico e cap)** e **un insieme di recapiti telefonici**. Ogni cliente è identificato da un codice che corrisponde al numero della tessera rilasciatagli per usufruire dei servizi della videoteca. **Ogni cliente è anche identificato dal proprio codice fiscale**. Ogni cliente può avere contemporaneamente in noleggio un certo numero di video (non più di tre). Per ogni noleggio vogliamo memorizzare **il giorno, mese ed anno** in cui il noleggio è stato effettuato e, per i noleggi conclusi, **il giorno, mese ed anno** della restituzione. Ogni cliente può inoltre consigliare dei film ad altri clienti, esprimendo per essi un giudizio **in una scala di valori interi da 0 a 5**. La videoteca prevede un programma di fidelizzazione dei clienti. Ogni noleggio consente l'accumulo di un certo numero di punti. Quando i punti accumulati superano una certa soglia, i clienti vengono qualificati come clienti VIP e hanno diritto ad un bonus **di un certo importo espresso in euro**. La base di dati dovrà memorizzare, per ogni cliente VIP, il valore corrente del suo bonus. Per i clienti standard, ovvero quelli che non hanno ancora accumulato punti sufficienti per **diventare clienti VIP**, vogliamo memorizzare il numero di punti mancanti a **diventarlo**.

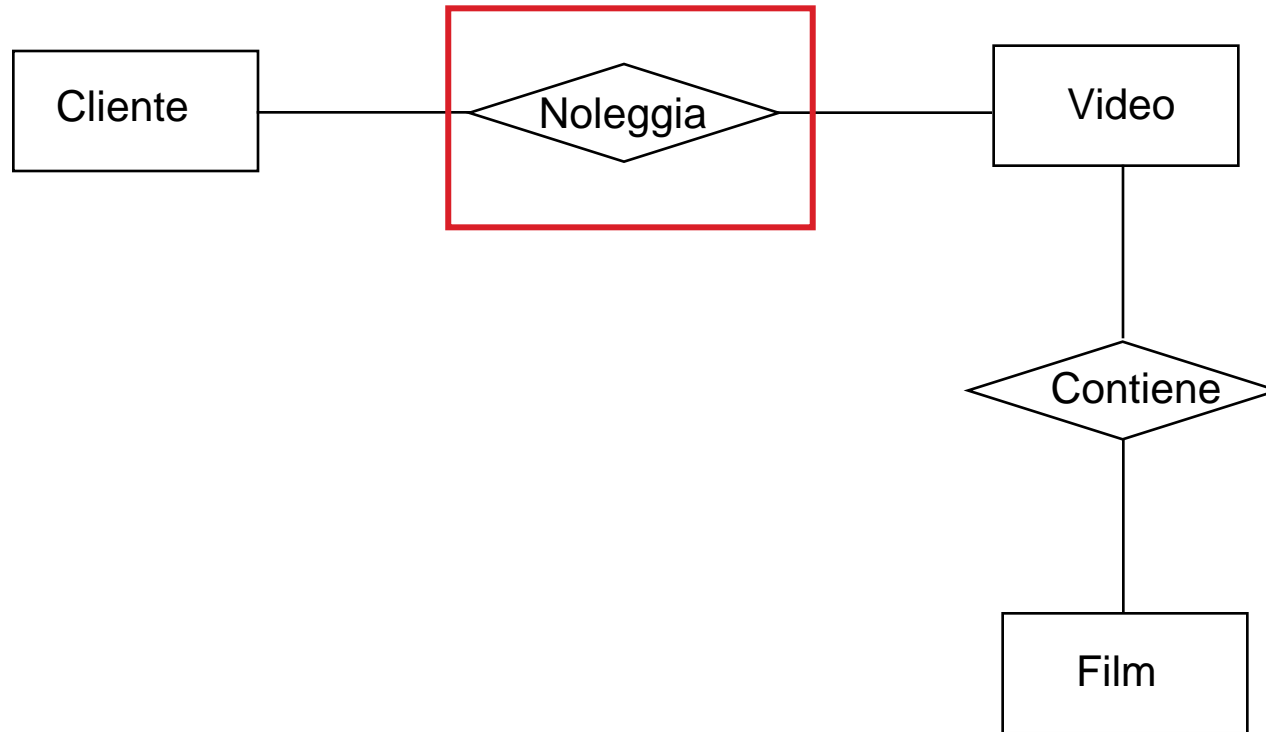
SCHEMA SCHELETRO



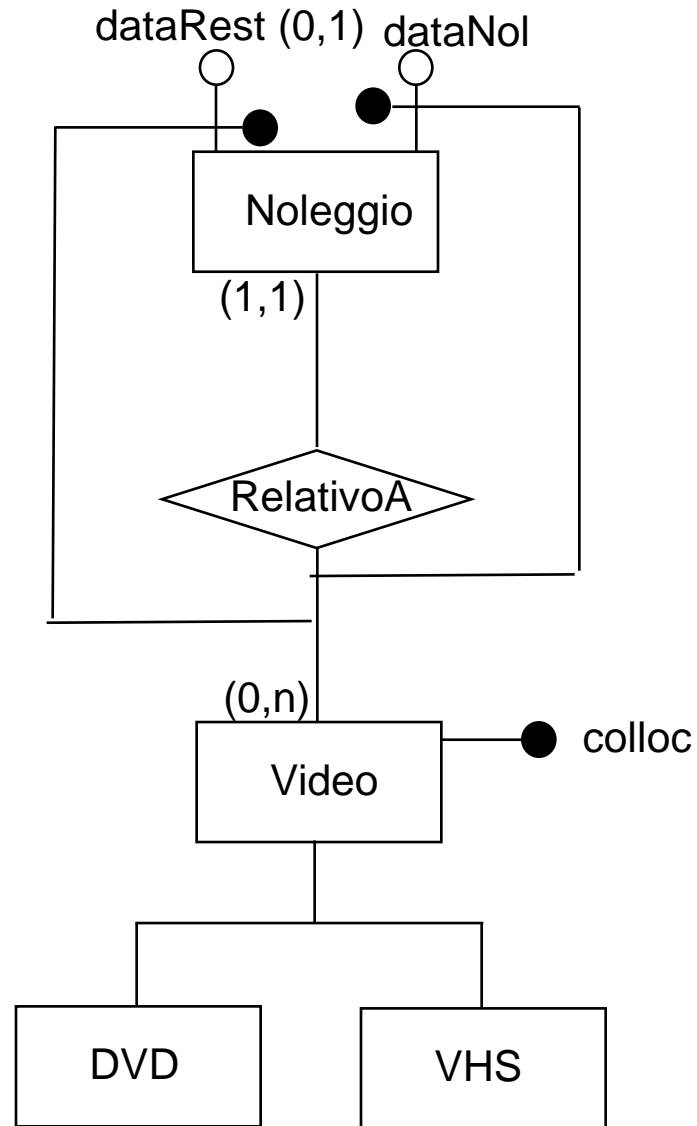
SOTTO-SCHEMA CLIENTE



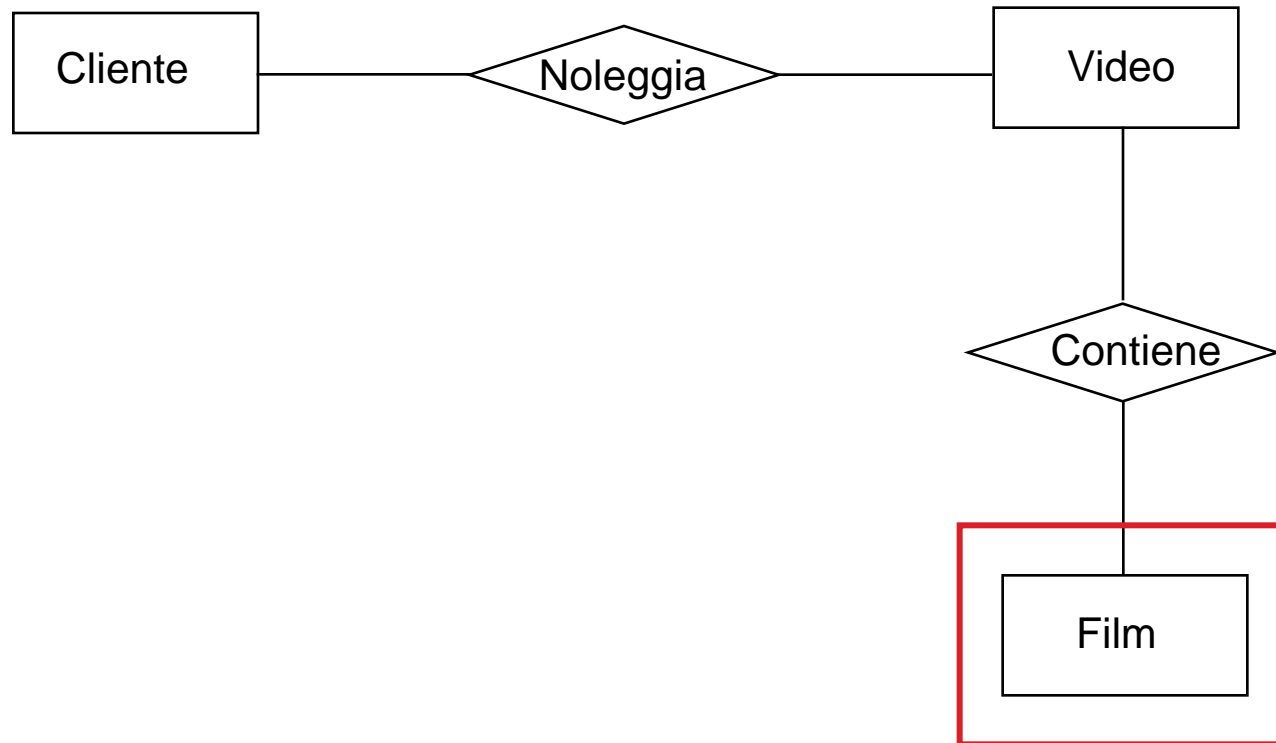
SCHEMA SCHELETRO



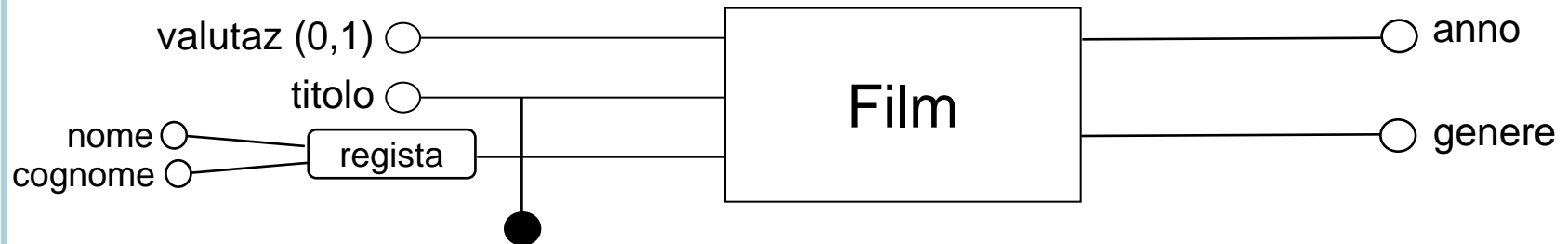
SOTTO-SCHEMA NOLEGGIO



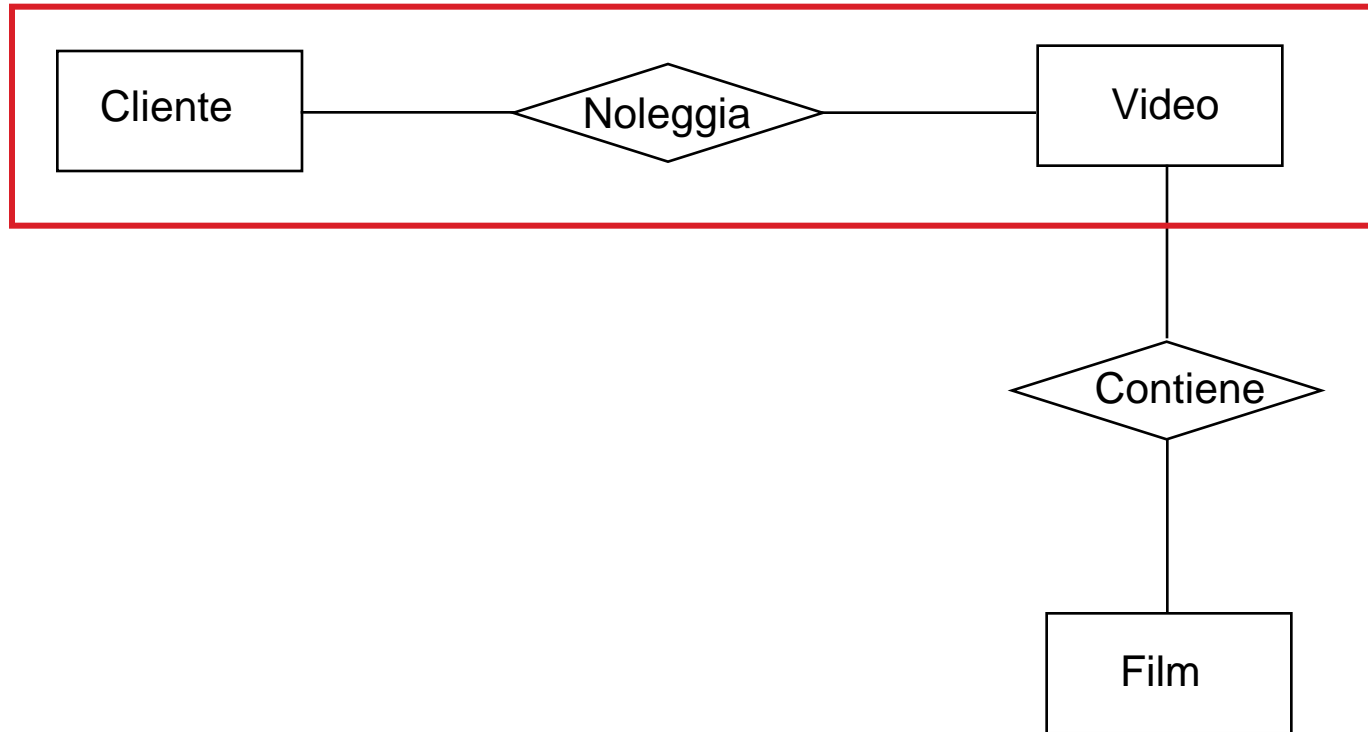
SCHEMA SCHELETRO



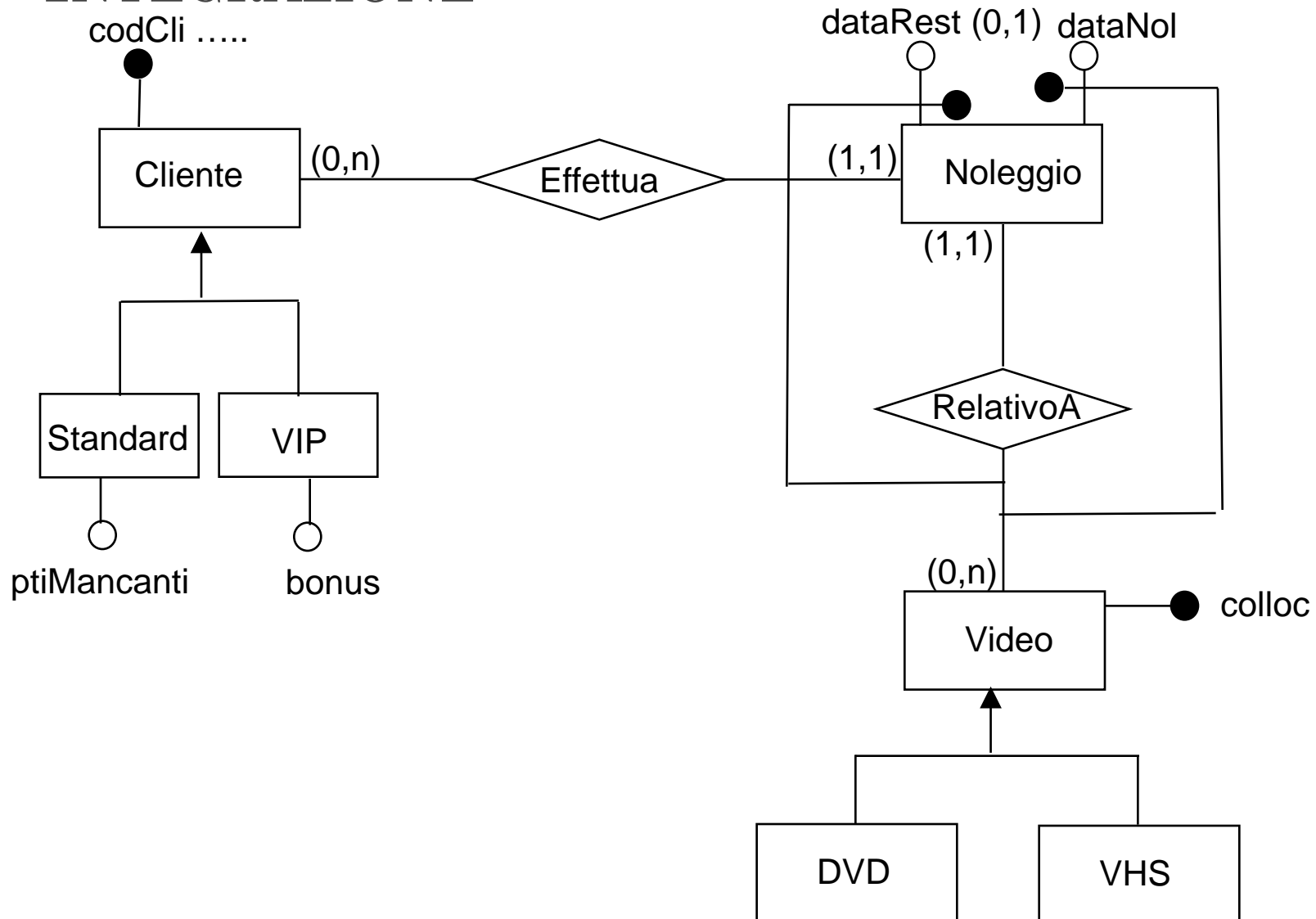
SOTTO-SCHEMA FILM



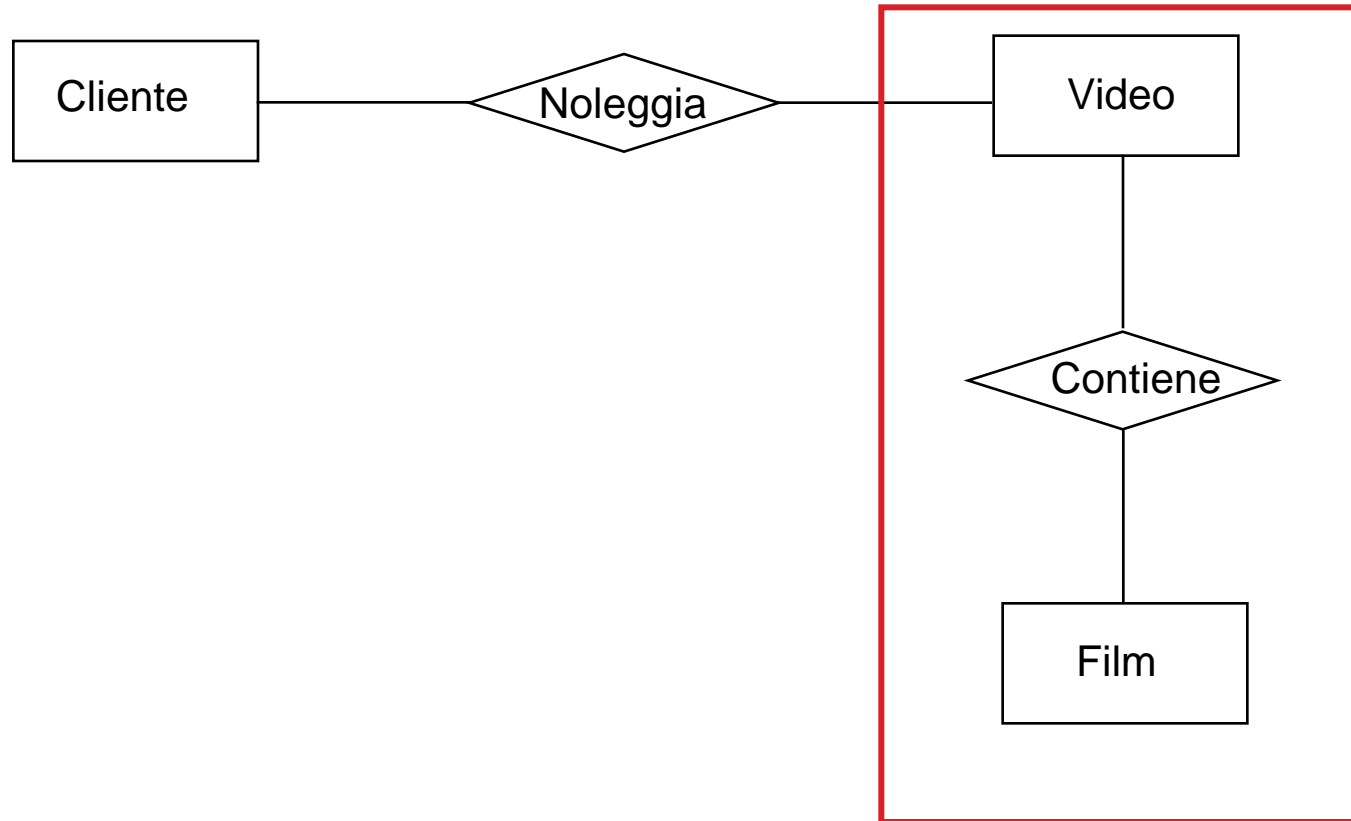
INTEGRAZIONE



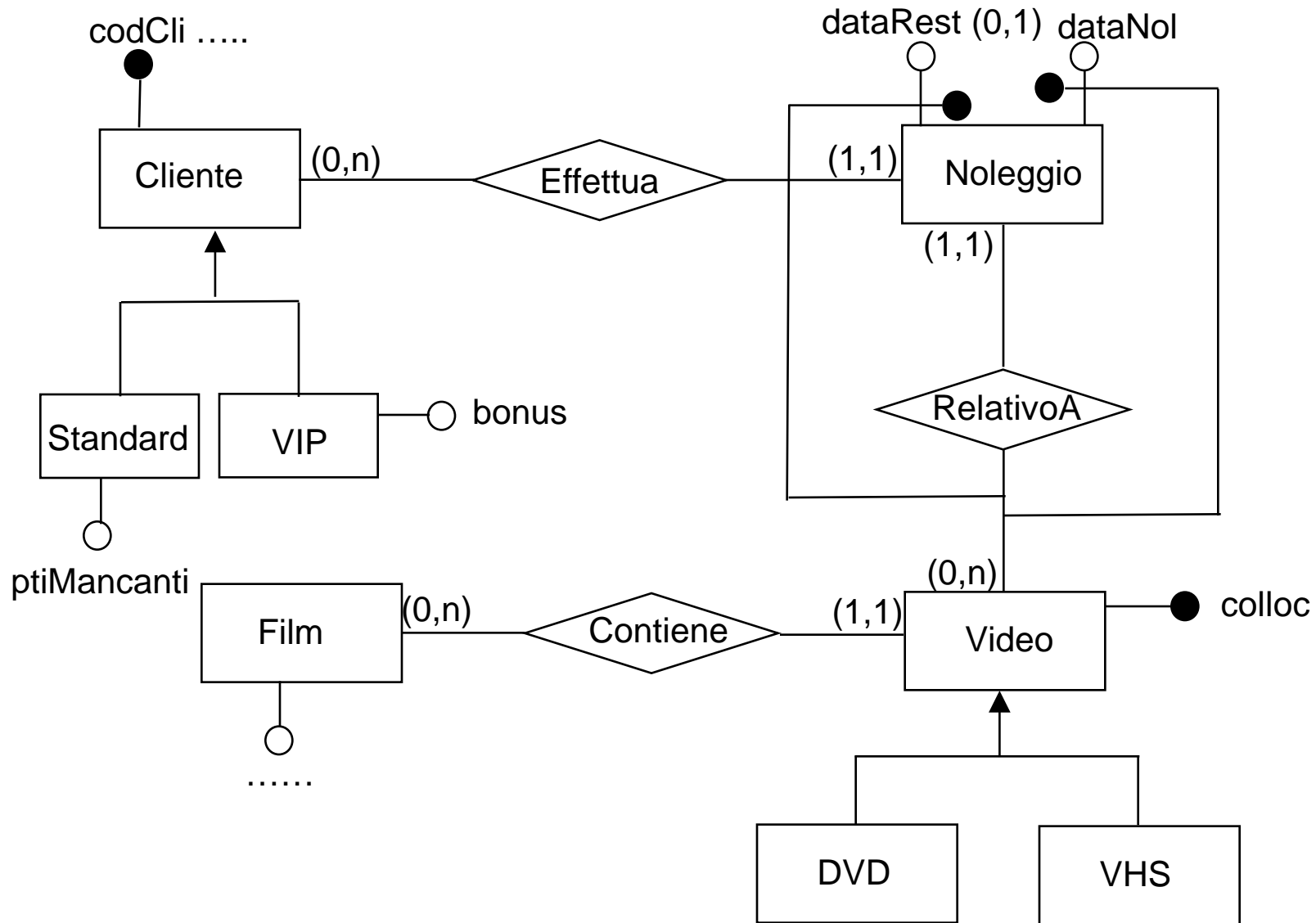
INTEGRAZIONE

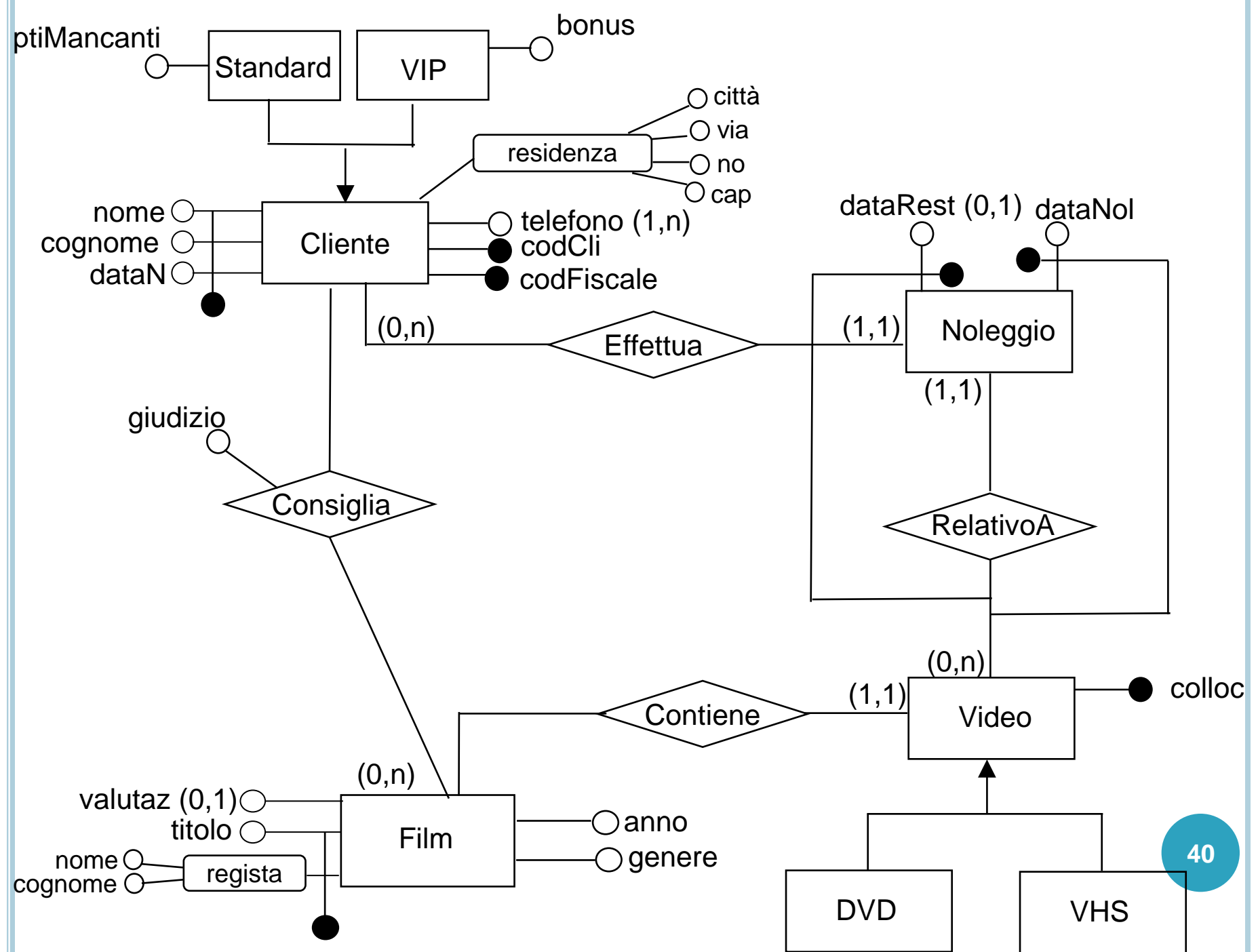


INTEGRAZIONE



INTEGRAZIONE





VINCOLI

V1: Un cliente non può noleggiare più di tre video contemporaneamente

V2: Un video non può essere noleggiato prima dell'uscita del relativo film

V3: La data di noleggio di un video non può essere successiva a quella di restituzione

V4: Uno stesso video non può essere noleggiato da due o più clienti diversi contemporaneamente

GERARCHIE DI GENERALIZZAZIONE

Entità padre	Entità figlie	Tipologia
Cliente	Standard, VIP	Totale/esclusiva
Video	DVD, VHS	Totale/esclusiva

DIZIONARIO DELLE ENTITÀ

Nome	Descrizione	Attributi	Identificatori
Cliente	Utenti che usufruiscono dei servizi della videoteca	codCli,dataN, nome, cognome, telefono, residenza, codFiscale	codCli, {nome, cognome, dataN}, codFiscale
Standard	Cliente comune della videoteca	ptiMancanti	gli stessi di Cliente
VIP	Cliente VIP della videoteca	bonus	gli stessi di Cliente
Video	Video offerti dalla videoteca	colloc	colloc
Noleggio	Noleggi correnti e conclusi effettuati nella videoteca	dataNol, dataRest	{video,dataNol} {video,dataRest}
Film	Film offerti dalla videoteca	titolo, regista, valutaz, anno, genere	{titolo, regista}
DVD	DVD offerti dalla videoteca		gli stessi di video
VHS	Videocassette offerte dalla videoteca		gli stessi di video

DIZIONARIO DELLE ASSOCIAZIONI

Nome	Descrizione	Attributi	Entità collegate
Effettuata	Noleggi effettuati		Cliente, Noleggio
RelativoA	Video noleggiati		Noleggio, Video
Contiene	Film a disposizione nella videoteca		Video, Film
Consiglia	Film consigliati dai clienti della videoteca	giudizio	Cliente, Film