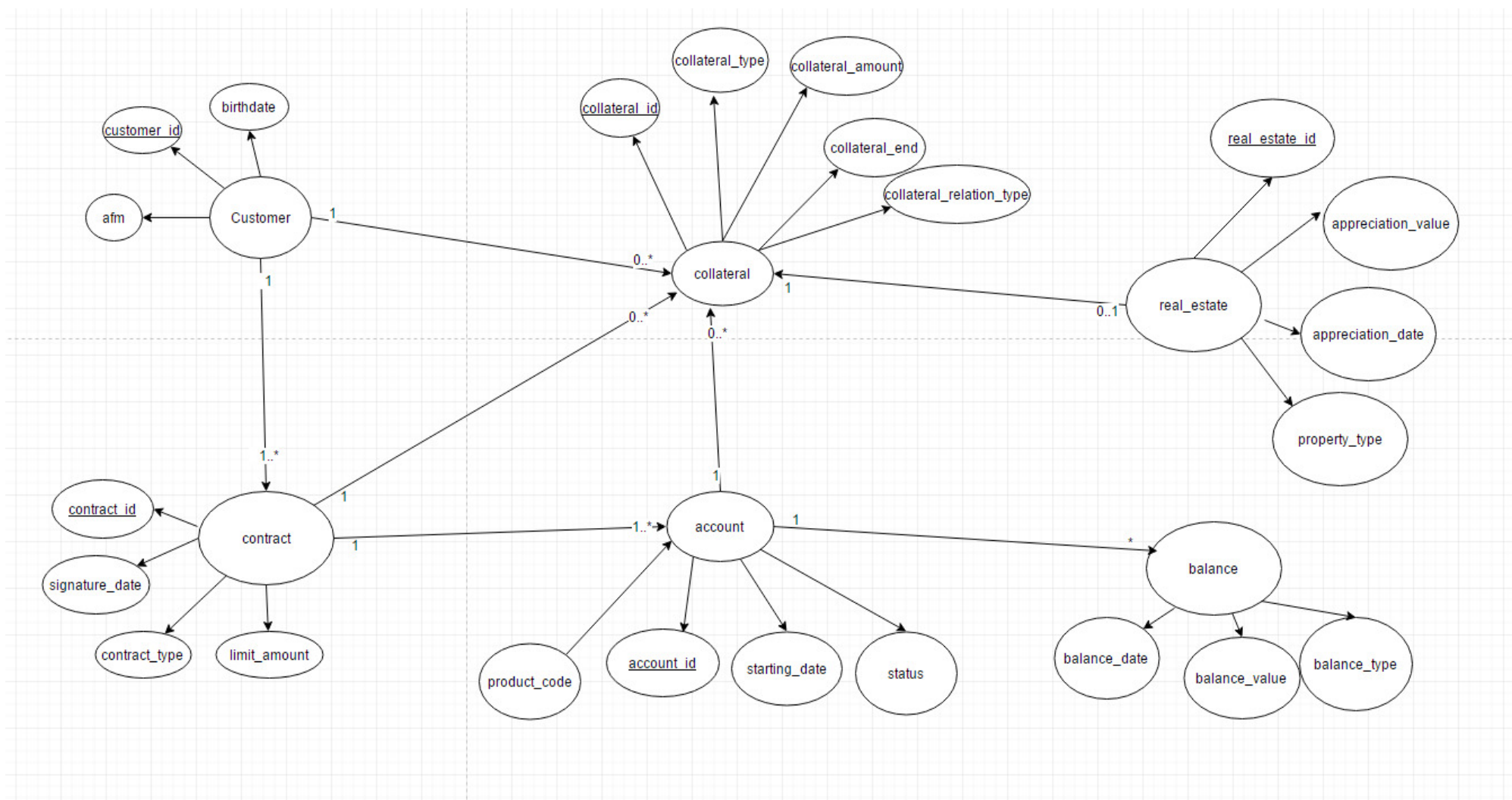


**ΣΧΕΔΙΑΣΜΟΣ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ**  
**PROJECT - ΜΕΡΟΣ Α**  
**ΟΝΟΜΑΤΕΠΩΝΥΜΟ: ΓΕΩΡΓΙΟΣ ΧΕΙΡΜΠΟΣ**  
**ΑΜ: 3130230**

Βάση τις διευκρινήσεις/επεξηγήσεις της εκφώνησης έβγαλα το παρακάτω σχημα E-R της βάσης.



### ... -> 1NF

Σχετικά με το dat (τον οποίο έχω ονομάσει ως Collateral [εξήγησεις στις περιοχές σχολίων περί κώδικα]) . Βρίσκεται ήδη στην 1η κανονική.

Πρωτεύον κλειδί είναι το collateral\_id. Αυτό επιβεβαιώνεται από το γεγονός ότι για δεδομένο collateral\_id θα πάρουμε μοναδική εγγραφή(row) από τον πίνακα. Επίσης σε κάθε στήλη υπάρχουν ατομικές τιμές, άρα δεν υπάρχει ανάγκη διασπαση.

Οι συναρτησιακές εξαρτήσεις που παρατηρούνται είναι οι εξής:

{customer\_id} -> birth\_date, afm

{contract\_id} -> signature\_date, limit\_amount, contract\_type, customer\_id

{account\_id} -> starting\_date, status, product\_code, contract\_id

{collateral\_id} -> collateral\_type, collateral\_amount, collateral\_end,

collateral\_relation\_type, account\_id, real\_estate\_id

{real\_estate\_id} -> appreciation\_value, appreciation\_date, property\_type

Για το bal (τον οποίο έχω ονομάσει ως Banance [εξήγησεις στις περιοχές σχολίων περί κώδικα]) για να προσδιορίσουμε μοναδικά κάθε εγγραφή στον πίνακα θα πρέπει να φτιάξουμε ένα σύνθετο πρωτεύον κλειδί το οποίο θα αποτελείτε από τις στήλες account\_id, balance\_date, balance\_value.

Αυτό συμβαίνει διότι για συγκεκριμένο λογαριασμό, υπάρχει συγκεκριμένο υπόλοιπο.

Αυτό το υπόλοιπο κάθε τέλος μια δεδομένης μέρα καταγράφεται η ημερομηνία, το είδος υπολοίπου και το ποσό του υπολοίπου.

Επόμενως ο συνδυασμός των λογαριασμού, ημερομηνίας, είδος υπολοίπου είναι μοναδικός για κάθε σειρά στον πίνακα. Οπότε με αυτή την ομαδοποίηση και δεδομένου ότι οι τιμές κάθε στήλης είναι ατομικές ο πίνακας βρίσκεται στο 1NF.

Οι συναρτησιακές εξαρτήσεις που παρατηρούνται είναι οι εξής:

{customer\_id} -> birth\_date, afm

{contract\_id} -> signature\_date, limit\_amount, contract\_type, customer\_id

{account\_id} -> starting\_date, status, product\_code, contract\_id

{account\_id, balance\_date, balance\_type} -> balance\_value

### 1NF -> 2NF

Το πρόβλημα εδώ βρίσκεται στον bal. Υπάρχουν γνωρίσματα τα οποία εξαρτώνται συναρτησιακά από μέρος του κλειδιού.

{account\_id} -> starting\_date, status, product\_code, contract\_id

{account\_id, balance\_date, balance\_type} -> balance\_value

Αναγκαστικά θα πρέπει να σπάσουμε τον πίνακα σε 4 κομμάτια. Αυτό για να μην χάσουμε τις εξαρτήσεις μεταξύ account με contract και contract με customer.

Οι πίνακες που θα σχηματιστούν θα είναι:

**Customer** με μορφή  
{customer\_id} -> birth\_date, afm

**Contract** με μορφή  
{contract\_id} -> signature\_date, limit\_amount, contract\_type, customer\_id  
όπου το customer\_id είναι foreign key και αναφέρετε στο customer\_id του Customer

**Account** με μορφή  
{account\_id} -> starting\_date, status, product\_code, contract\_id  
όπου το contract\_id είναι foreign key και αναφέρετε στο contract\_id του Contract

**Balance** με μορφή  
{account\_id, balance\_date, balance\_type} -> balance\_value  
Εδώ το account\_id θα παίξει το ρολό και του ξένου κλειδιού ώστε να μην χαθεί η σύνδεση του με τον πίνακα Account και θα παραμείνει κομμάτι του πρωτεύοντος κλειδιού.

Έτσι πετύχαμε το 2NF για αυτούς τους πίνακες.  
Ο dat βρίσκεται ήδη σε 2NF.

## 2NF -> 3NF

Εδώ οι νέοι πίνακες Customer, Contract, Account\_, Balance που δημιουργήθηκαν κατά την μετάβαση σε 2NF, είναι ήδη στην 3NF αφού δεν υπάρχουν εσωτερικές εξαρτήσεις μεταξύ των γνωρισμάτων παρά μόνο με το πρωτεύον κλειδί.

Όμως ο πίνακας dat περιέχει μεταβατικές εξαρτήσεις  
collateral\_id -> real\_estate\_id -> ...  
collateral\_id -> account\_id -> contract\_id -> customer\_id -> ...

Επομένως πρέπει να γίνει διάσπαση.  
Όσον αφορά τα Customer, Contract και Account\_ είναι ήδη δημιουργημένοι.  
Μένει να φτιάξουμε έναν πίνακα  
Real\_estate μορφή  
{real\_estate\_id} -> appreciation\_value, appreciation\_date, property\_type

και μένει ο Collateral με μορφή  
{collateral\_id} -> collateral\_type, collateral\_amount, collateral\_end,  
collateral\_relation\_type, account\_id, real\_estate\_id

Όπου το account\_id ως ξένο κλειδί αναφέρεται στον Account\_ και το real\_estate\_id στον Real\_estate.

Έτσι πετύχαμε το 3NF για τον πρώην dat.

### **3NF -> BCNF**

Από ότι παρατηρούμε όλοι οι πίνακες που έχουμε βρίσκονται στην BCNF μορφή αφού δεν σε κάθε πίνακα το primary key αποτελεί και υπερκλειδί κάθε συναρτησιακής εξάρτησης.

### **[Σχόλια περί κώδικα SQL]**

Τα αρχεία dat.txt και bal.txt φορτώνονται στους πίνακες Collateral και Balance αντίστοιχα. Αφού γίνουν τα create των Customer, Contract, Account\_, Real\_estate φορτώνονται τα δεδομένα.

Οι Customer, Contract, Account\_, Real\_estate πέρνουν τα δεδομένα από αυτούς.

Για λόγους απόδοσης οι αρχικοί "προσωρινοί" πίνακες περνάνε από διαδικασία alter ώστε να τροποποιηθούν κατάλληλα για τους άλλους δύο πίνακες που χρειαζόμαστε.

Αυτό γίνεται κάνοντας drop τα column που δεν χρειαζόμαστε και ορίζοντας τα αντίστοιχα ξένα και πρωτεύονται κλειδιά.

#### **#Query 1**

Χρειαζόμαστε μόνο δυο πίνακες όπου ελέγχοντας με ακρίβεια διαφορά ετών(signature\_date - birthdate) μετράμε διακριτά όσοι είναι κάτω από 19 όταν υπέγραψαν σύμβαση.

#### **#Query 2**

Χρησιμοποιώντας την ίδια ακρίβεια όπως στο προηγούμενο ερώτημα (αυτή την φορά "τωρινή ημερομηνία" – birthdate) βρίσκουμε διακριτά όσοι είναι πάνω από 60 που καλύπτονται με ακίνητο στην πρώτη κατοικία άμεσα.

#### **#Query 3**

Εδώ απαιτείται η χρήση subquery για να γίνει σωστή αντιστοίχιση την πιο πρόσφατης ημερομηνίας ανα λογαριασμό και υπολοίπου κεφαλαίο και υπόλοιπο κεφαλαίου μεγαλύτερο του 1.000.000. Έπειτα εντοπίζονται και μετριοούνται διακριτά όσοι πληρούν τις προϋποθέσεις.

#### **#Query 4**

Χρησιμοποιώντας τον πίνακα Collateral μετράμε τον αριθμό των collateral και το αξία τους και τα ομαδοποιούμε ανά λογαριασμό.