



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

**Απόδοση Πληροφορίας Αποθηκευμένης σε Βάσεις
Δεδομένων με Φυσική Γλώσσα**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΙΩΑΝΝΗ ΑΛΕΞΑΝΔΡΑΚΗ

Επιβλέπων : Τιμολέων Σελλής
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2007



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Απόδοση Πληροφορίας Αποθηκευμένης σε Βάσεις Δεδομένων με Φυσική Γλώσσα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΙΩΑΝΝΗ ΑΛΕΞΑΝΔΡΑΚΗ

Επιβλέπων : Τιμολέων Σελλής
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 8^η Οκτωβρίου 2007.

.....
Τιμολέων Σελλής
Καθηγητής Ε.Μ.Π.

.....
Ιωάννης Βασιλείου
Καθηγητής Ε.Μ.Π.

.....
Νεκτάριος Κοζύρης
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2007

.....
ΙΩΑΝΝΗΣ ΑΛΕΞΑΝΔΡΑΚΗΣ

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Ιωάννης Αλεξανδράκης, 2007

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Οι χρήστες χωρίς γνώση σχημάτων ή γλωσσών διατύπωσης ερωτήσεων έχουν δυσκολίες στην πρόσβαση πληροφοριών που αποθηκεύονται στις βάσεις δεδομένων. Οι εμπορικές και ερευνητικές προσπάθειες έχουν εστιάσει στις αναζητήσεις με λέξεις κλειδιά. Μεταξύ αυτών των εργασιών, οι *precis queries* παράγουν ολόκληρες πολυ-σχεσιακές βάσεις δεδομένων, που είναι λογικά υποσύνολα των υπαρχόντων, αντί των μεμονωμένων σχέσεων.

Το λογικό υποσύνολο της βάσης δεδομένων περιέχει όχι μόνο στοιχεία σχετικά άμεσα με τις επιλογές ερώτησης αλλά και λοιπά στοιχεία σχετικά με αυτές. Προηγούμενη δουλειά έχει προσδιορίσει την ανάγκη να δώσει στον απλό χρήστη απαντήσεις με νόημα στις ερωτήσεις του και έχει προτείνει τη μετάφραση των *precis queries* με αφηγηματική μορφή.

Σε αυτήν την διπλωματική εργασία, υλοποιήθηκε μια ημιαυτόματη μέθοδος που μεταφράζει το παραγόμενο σύνολο σχέσεων (μια βάση δεδομένων – υποσύνολο της αρχικής) μιας ερώτησης *precis* σε μια σύνθεση αποτελεσμάτων. Περιγράφουμε την μηχανή μετάφρασης που δημιουργήθηκε και που χρησιμοποιεί έναν μηχανισμό προτύπων για την παραγωγή μιας ερώτησης *precis* σε αφηγηματική μορφή μέσω ενός συνόλου επαναχρησιμοποιήσιμων προτύπων. Περιγράφουμε επίσης ένα φιλικό προς το χρήστη γραφικό περιβάλλον που αναπτύχθηκε, προκειμένου να βοηθήσουμε έναν χρήστη να δημιουργήσει τα αναγκαία προαναφερθέντα πρότυπα και να επιβλέπει ολόκληρη τη διαδικασία.

Λέξεις Κλειδιά: Αναζήτηση με λέξεις-κλειδιά, βελτιστοποίηση επερωτήσεων, πρότυπα, φυσική γλώσσα, *precis queries*.

Abstract

Users without knowledge of schemas or query languages have difficulties in accessing information stored in databases. Commercial and research efforts have focused on keyword-based searches. Among them, *précis* queries generate entire multi-relation databases, which are logical subsets of existing ones, instead of individual relations.

The logical database subset contains not only items directly related to the query selections but also items implicitly related to them in various ways. Earlier work has identified the need of providing the naïve user with meaningful answers to his questions and has suggested the translation of *précis* query answer in narrative form.

In this thesis, we implement a semi-automatic method that translates the relational output of a *précis* query into a synthesis of results. We describe the translator engine that was implemented and which uses a template mechanism for generating a *précis* in a narrative form through a set of reusable templates. We also describe a user-friendly, Graphical User Interface that was implemented, in order to help a user create the needed templates mentioned above and overview the whole process.

Keywords: Keyword search, query optimization, templates, natural language, *précis* queries.

Πίνακας περιεχομένων

| | | |
|----------|---|-----------|
| 1 | Εισαγωγή | 1 |
| 1.1 | Εξαγωγή λογικών απαντήσεων από δομημένα δεδομένα | 1 |
| 1.2 | Αντικείμενο διπλωματικής | 2 |
| 1.2.1 | <i>Συνεισφορά</i> | 3 |
| 1.3 | Οργάνωση κειμένου | 3 |
| 2 | Σχετικές εργασίες | 5 |
| 3 | Θεωρητικό υπόβαθρο | 9 |
| 3.1 | Εισαγωγικά | 9 |
| 3.2 | Μοντέλο ερωτήσεων précis | 12 |
| 3.3 | Λογική μετάφραση υποσυνόλων | 13 |
| 3.3.1 | <i>Προκαταρκτικά</i> | 13 |
| 3.3.2 | <i>Μηχανισμός προτύπων</i> | 16 |
| 3.3.3 | <i>Μετάφραση</i> | 21 |
| 3.3.4 | <i>Δυνατότητα εφαρμογής και περιορισμοί</i> | 29 |
| 4 | Σχεδίαση Συστήματος | 31 |
| 4.1 | Αρχιτεκτονική | 33 |
| 4.1.1 | <i>Μηχανή μετάφρασης</i> | 33 |
| 4.1.2 | <i>Γραφική διεπαφή χρήστη (GUI)</i> | 35 |
| 4.2 | Περιγραφή Κλάσεων | 39 |
| 4.2.1 | <i>Μηχανή μετάφρασης</i> | 39 |
| 4.2.2 | <i>Γραφική διεπαφή χρήστη</i> | 42 |
| 4.3 | Βάσεις δεδομένων | 56 |
| 5 | Υλοποίηση | 61 |
| 5.1 | Λεπτομέρειες υλοποίησης | 61 |
| 5.1.1 | <i>Υλοποίηση γράφου (κόμβων και ακμών)</i> | 61 |
| 5.1.2 | <i>Algorithm Resolve Common Expressions</i> | 67 |
| 5.1.3 | <i>Algorithm Translation of a Logical Subset (TRLS)</i> | 70 |

| | | |
|----------|---|------------|
| 5.1.4 | Algorithm Translation of a Relation (TR)..... | 71 |
| 5.1.5 | Algorithm Unary Translation (UTR)..... | 74 |
| 5.1.6 | Algorithm Split Translation (STR)..... | 75 |
| 5.1.7 | Algorithm Translation of a macro..... | 77 |
| 5.2 | Πλατφόρμες και προγραμματιστικά εργαλεία..... | 79 |
| 5.2.1 | Λειτουργικό σύστημα..... | 79 |
| 5.2.2 | Γλώσσες προγραμματισμού..... | 79 |
| 5.2.3 | Σύστημα Διαχείρισης Βάσεων Δεδομένων (ORACLE)..... | 82 |
| 6 | Έλεγχος..... | 85 |
| 6.1 | Τρόπος χρήσης..... | 85 |
| 6.1.1 | Γραφική διεπαφή χρήστη..... | 85 |
| 6.2 | Σενάρια – αποτελέσματα μετάφρασης..... | 98 |
| 6.3 | Σύνοψη και επέκταση αποτελεσμάτων..... | 137 |
| 7 | Επίλογος..... | 141 |
| 7.1 | Σύνοψη και συμπεράσματα..... | 141 |
| 7.2 | Μελλοντικές επεκτάσεις..... | 144 |
| 8 | Βιβλιογραφία..... | 145 |
| | ΠΑΡΑΡΤΗΜΑ Α..... | 147 |
| | Μετάφραση χωρίς περιορισμούς..... | 147 |
| | Όριο πέντε μεταφραζόμενων πλειάδων στις μακροεντολές..... | 158 |
| | Όριο τριών μεταφραζόμενων πλειάδων στις μακροεντολές..... | 164 |
| | Όριο μιας μεταφραζόμενης πλειάδας στις μακροεντολές..... | 169 |

1

Εισαγωγή

1.1 Εξαγωγή λογικών απαντήσεων από δομημένα δεδομένα

Η ανάγκη για πρόσβαση στις πληροφορίες που αποθηκεύονται στις βάσεις δεδομένων έχει αναγνωριστεί από νωρίς από την ερευνητική κοινότητα με αρχικές προσπάθειες που χρονολογούνται από τη δεκαετία του '70 [7]. Η εμφάνιση του Διαδικτύου (World Wide Web) έχει καταστήσει δυνατή την πρόσβαση πληροφοριών σε έναν αυξανόμενο αριθμό ανθρώπων. Ένα μεγάλο μέρος των πληροφοριών ενυπάρχει στις βάσεις δεδομένων, δεδομένου ότι οι βιβλιοθήκες, τα μουσεία, και άλλες οργανώσεις δημοσιεύουν το ηλεκτρονικό περιεχόμενό τους στον Παγκόσμιο Ιστό. Παράλληλα, οι περισσότεροι χρήστες δεν έχουν καμία συγκεκριμένη γνώση σχημάτων ή δομημένων γλωσσών διατύπωσης ερωτήσεων (structured query languages) για την πρόσβαση των πληροφοριών που αποθηκεύονται σε μια βάση δεδομένων. Σε αυτό το πλαίσιο, η ανάγκη για την πρόσβαση σε αυτές τις πληροφορίες που αποθηκεύονται στις βάσεις δεδομένων γίνεται όλο και περισσότερο σημαντικότερη. Οι υπάρχουσες προσπάθειες έχουν εστιάσει κυρίως στη δημιουργία ερωτήσεων (queries) στις σχεσιακές βάσεις δεδομένων προτείνοντας είτε διαχείριση ερωτήσεων φυσικής γλώσσας [2, 13, 16] είτε ελεύθερης-μορφής, βασισμένες σε λέξεις-κλειδιά, ερωτήσεις [1, 18].

1.2 Αντικείμενο διπλωματικής

Σε αυτήν την διπλωματική εργασία, εστιάζουμε σε ένα σχετικό, πρωτότυπο, σύστημα παραγωγής απαντήσεων στις ερωτήσεις και προτείνουμε μια προσέγγιση για να μεταφράσουμε τη σχεσιακή παραγωγή μιας ερώτησης σε μια μορφή που μοιάζει με την αφήγηση και είναι έτσι πιο κατανοητή στον απλό χρήστη.

Ειδικότερα, εξετάζουμε την παραγωγή των ερωτήσεων *précis* (*précis queries*) [11]. Αυτές είναι ελεύθερης-μορφής ερωτήσεις που *queries* παράγουν ολόκληρες πολυ-σχεσιακές βάσεις δεδομένων, που είναι λογικά υποσύνολα των υπάρχοντων, αντί μεμονωμένων σχέσεων. Το λογικό υποσύνολο της βάσης δεδομένων περιέχει όχι μόνο τα στοιχεία που είναι άμεσα σχετικά με τις επιλογές ερώτησης αλλά και λοιπά στοιχεία έμμεσα σχετιζόμενα με αυτές. Τα λογικά υποσύνολα βάσεων δεδομένων είναι χρήσιμα σε πολλές περιπτώσεις. Λαμβάνοντας υπόψη τις μεγάλες βάσεις δεδομένων, οι επιχειρήσεις έχουν ανάγκη συχνά τα μικρότερα υποσύνολα τους που προσαρμόζονται στο αρχικό σχήμα και ικανοποιούν κάποιους περιορισμούς προκειμένου να εκτελέσουν ρεαλιστικές δοκιμές νέων εφαρμογών πριν επεκταθούν στην παραγωγή. Επίσης, οι προμηθευτές λογισμικού χρειάζονται τέτοιες μικρότερες αλλά ορθά δομημένες βάσεις δεδομένων για να καταδείξουν τις λειτουργίες των νέων προϊόντων λογισμικού. Επιπλέον, οι απλοί χρήστες θα ανέμεναν μάλλον μια περίληψη ή σύνοψη (*précis*) των πληροφοριών που περιλαμβάνονται σε ένα λογικό υποσύνολο. Παραδείγματος χάριν, μια πιο ορθή νοηματικά απόκριση από την κλασική απάντηση που είναι σε μορφή πίνακα σε μια ερώτηση (*query*) που ρωτά για τον “Woody Allen” να είναι υπό μορφή ακόλουθων αποτελεσμάτων *précis* ερωτήσεων :

“Woody Allen was born on December 1, 1935 in Brooklyn, New York, USA.

As a director, Woody Allen’s work includes Match Point (2005), Melinda and Melinda (2004), Anything Else (2003). As an actor, Woody Allen’s work includes Hollywood Ending (2002), The Curse of the Jade Scorpion (2001).”

Οι *précis* μπορεί να είναι ελλιπείς από πολλές απόψεις, παραδείγματος χάριν, οι προαναφερθέντες *précis* περιλαμβάνουν έναν μη λεπτομερή κατάλογο έργων του Woody Allen. Εντούτοις, παρέχει ικανοποιητικές πληροφορίες με έναν κατανοητό τρόπο για να βοηθήσει κάποιον να μάθει για τον Woody Allen και να προσδιορίσει ενδεχομένως νέες λέξεις κλειδιά για περαιτέρω αναζήτηση. Παραδείγματος χάριν, κάποιος μπορεί να αποφασίσει να δημιουργήσει μια νέα ερώτηση για την ταινία “*Anything Else (2003)*”.

1.2.1 Συνεισφορά

Αυτή η εργασία υλοποιεί την παρουσίαση μιας απάντησης *précis* σε μια ερώτηση βασισμένη σε λέξεις-κλειδιά σε μια σχεσιακή βάση δεδομένων. Εν συντομία :

- Περιγράφουμε τη λειτουργία των ερωτήσεων *précis*, εμπλουτίζοντας το πρότυπο με ετικέτες (labels) συνδεδεμένες στα επιμέρους μέρη της (σχέσεις, πεδία). Επίσης, παρουσιάζουμε έναν απλό τρόπο να συνθέσουμε αυτές τις ετικέτες μέσω μιας απλής στην χρήση γλώσσας.
- Παρουσιάζουμε έναν μηχανισμό για τον καθορισμό και την αρχικοποίηση των προτύπων ετικετών.
- Παρουσιάζουμε μια ημιαντόματη μέθοδο που μεταφράζει το παραγόμενο σύνολο σχέσεων (μια βάση δεδομένων – υποσύνολο της αρχικής) μιας ερώτησης *précis* σε μια σύνθεση αποτελεσμάτων.

1.3 Οργάνωση κειμένου

Εργασίες σχετικές με το αντικείμενο της διπλωματικής παρουσιάζονται στο Κεφάλαιο 2. Στο κεφάλαιο 3 αναλύεται το θεωρητικό υπόβαθρο της εργασίας. Το κεφάλαιο 4 αφορά στην σχεδίαση του συστήματος, την αρχιτεκτονική των δυο εργαλείων που αναπτύχθηκαν, της μηχανής μετάφρασης και της γραφικής διεπαφής χρήστη. Στο 5^ο κεφάλαιο παρουσιάζονται λεπτομέρειες υλοποίησης των εργαλείων καθώς και αναλύονται αλγοριθμικά μοντέλα που χρησιμοποιήθηκαν. Στο κεφάλαιο 6 παρουσιάζονται σενάρια χρήσης των εργαλείων μαζί με τα αποτελέσματά τους, καθώς και ο τρόπος χρήσης των εργαλείων. Στο 7^ο κεφάλαιο βρίσκονται τα συμπεράσματα που εξήχθησαν από τα αποτελέσματα και συζητούνται μελλοντικές επεκτάσεις των εργαλείων και τέλος στο 8^ο κεφάλαιο βρίσκεται η χρησιμοποιούμενη βιβλιογραφία για την παρούσα εργασία.

2

Σχετικές εργασίες

Η ανάγκη για δημιουργία ερωτήσεων ελεύθερης-μορφής έχει αναγνωριστεί από νωρίς στα πλαίσια των βάσεων δεδομένων [18]. Με την εμφάνιση του Παγκόσμιου Ιστού (World Wide Web), η ιδέα έχει έρθει ξανά στην επιφάνεια. Διάφορες ερευνητικές προσπάθειες έχουν προκύψει για λέξεις – κλειδιά σε δομημένα [1, 3, 8, 12] ή ημι-δομημένα [5, 6, 9] δεδομένα. Τα Oracle 9i Text [19], Microsoft SQL Server [15] and IBM DB2 Text Information Extender [10] δημιουργούν πλήρεις δείκτες κειμένων (text indexes) σε πεδία κειμένου των σχέσεων και πραγματοποιούν έπειτα τις αναζητήσεις με λέξεις κλειδιά.

Οι υπάρχουσες αναζητήσεις με λέξεις-κλειδιά εστιάζουν στην εύρεση και ενδεχομένως διασύνδεση πλειάδων στις σχέσεις που περιέχουν τους όρους ερώτησης. Παραδείγματος χάριν, η απάντηση για τον “Woody Allen” θα ήταν υπό μορφή ζευγαριού σχέσης-πεδίων, όπως (ΣΚΗΝΟΘΕΤΗΣ, ΟΝΟΜΑ). Σε πολλές περιπτώσεις, αυτή η απάντηση μπορεί να αρκεί, αλλά σε πολλά πρακτικά σενάρια περιέχει λιγοστές πληροφορίες για τον “Woody Allen”. Μια πληρέστερη απάντηση που περιέχει, παραδείγματος χάριν, τις πληροφορίες για τους κινηματογράφους και τα βραβεία αυτού του σκηνοθέτη θα ήταν σημαντικότερη και πιο χρήσιμη. Ακολουθώντας αυτό το πνεύμα, πρόσφατα, οι ερωτήσεις *précis* έχουν προταθεί [11]. Αυτές είναι ερωτήσεις ελεύθερης-μορφής οι οποίες αντί να εντοπίζουν και απλά να συνδέουν τιμές στους πίνακες, θεωρούν επίσης ότι πληροφορίες σχετικές με αυτές τις τιμές μπορεί να τις αφορούν. Επομένως, η απάντηση σε μια ερώτηση *précis* επίσης να περιέχει πληροφορίες που βρέθηκαν σε άλλα μέρη της βάσης δεδομένων, π.χ., ταινίες που σκηνοθετήθηκαν από τον Woody Allen. Αυτές οι πληροφορίες πρέπει να συγκεντρωθούν – πιθανόν με απρόβλεπτους τρόπους - με την ένωση πλειάδων από πολλαπλές σχέσεις. Συνεπώς, η απάντηση σε μια ερώτηση *précis* είναι μια νέα βάση δεδομένων που είναι λογικό υποσύνολο της αρχικής, σε αντιδιαστολή με τα μονοδιάστατα αποτελέσματα των άλλων προσεγγίσεων.

Τα λογικά υποσύνολα βάσεων δεδομένων είναι χρήσιμα σε πολλές περιπτώσεις. Εντούτοις, οι απλοί χρήστες θα προτιμούσαν μάλλον μια φιλική αντιπροσώπευση των πληροφοριών που περιλαμβάνονται σε ένα λογικό υποσύνολο, χωρίς απαραίτητα να καταλάβουν τον σχεσιακό χαρακτήρα του. Σε άλλη εργασία [11], η σημασία τέτοιας εμφάνισης που κατασκευάζεται βασισμένης στις πληροφορίες που υπάρχουν στην αναπαράσταση σε γράφο μιας βάσης δεδομένων, έχει εμφανιστεί. Αυτό εμπνέεται από το BAROQUE [17] και προστατεύει το χρήστη από τις ιδιαιτερότητες του υπό εξέταση σχήματος και του μοντέλου δεδομένων που χρησιμοποιείται. Το BAROQUE χρησιμοποιεί μια δικτυακή εμφάνιση μιας βάσης δεδομένων και ορίζει αρκετούς τύπους σχέσεων προκειμένου να υποστηριχθούν λειτουργίες που ανιχνεύουν αυτό το δίκτυο. Εντούτοις, εντοπίζει μόνο τη θέση και τις σχέσεις στις οποίες ένα στοιχείο συμμετέχει. Δεδομένου ότι η αντιπροσώπευση βάσεων δεδομένων που υιοθετείται δεν περιλαμβάνει ενώσεις, δεν μπορεί να συγκεντρώσει τις απαντήσεις που χωρίζονται σε διάφορες σχέσεις ενώνοντάς τες.

Το πρόβλημα του απλού χρήστη έχει συζητηθεί λεπτομερώς στον τομέα της επεξεργασίας φυσικής γλώσσας (NLP). Τις τελευταίες δεκαετίες, διάφορες εργασίες έχουν παρουσιαστεί σχετικά με NL Querying [26, 14], NL and Schema Design [22, 13, 4], NL and DB interfaces [16, 2], and Question Answering [24, 21]. Όσο γνωρίζουμε, η σχετική βιβλιογραφία σχετικά με NL και τις βάσεις δεδομένων, έχει εστιάσει σε τελείως διαφορετικά ζητήματα όπως η ερμηνεία των φραστικών ερωτήσεων των χρηστών σε μια γλώσσα βάσεων δεδομένων, π.χ., SQL, ή στην αυτόματη σχεδίαση βάσεων δεδομένων, όπως για παράδειγμα με τη χρήση οντολογιών [23]. Υπάρχουν μερικές πρόσφατες προσπάθειες που χρησιμοποιούν τα φραστικά υποδείγματα ή πρότυπα ερώτησης για να διευκολύνουν τη διαδικασία απάντησης [16, 21]. Επίσης, υπάρχει μια πρόσφατη πειραματική μελέτη [26] που συγκρίνει τη NL Querying έναντι της αναζήτησης βασισμένης σε λέξεις κλειδιά και τάσσεται υπέρ της χρησιμότητας της τελευταίας ειδικά όταν αφορά σύνθετες ερωτήσεις.

Αυτή η εργασία βασίζεται στην παραγωγή απαντήσεων με νόημα από ερωτήσεις βασισμένες σε λέξεις - κλειδιά και αναπτύσσει μια προσέγγιση για να μεταφράσει το παραγόμενο σύνολο σχέσεων μιας ερώτησης σε αφηγηματική μορφή για να είναι έτσι πιο κατανοητή σε έναν απλό χρήστη. Η διαδικασία έχει κοινά γνωρίσματα με εκείνες που εμπεριέχουν χειρισμό μιας ερώτησης φυσικής γλώσσας σε σχεσιακές βάσεις δεδομένων στον βαθμό που και οι δυο περιλαμβάνουν κάποιους επιπρόσθετους ορισμούς για τις έννοιες που αντιπροσωπεύουν οι σχέσεις, οι ιδιότητες και ενώσεις πρωτευόντων – ξένων κλειδιών. Ωστόσο, η επεξεργασία ερώτησης φυσικής γλώσσας είναι πιο σύνθετη, δεδομένου ότι πρέπει να χειριστεί τις ασάφειες στη σύνταξη και τη σημασιολογία φυσικής γλώσσας ενώ η προσέγγισή μας χρησιμοποιεί ορθά καθορισμένα πρότυπα για να διαμορφώσει εκ νέου τις σχέσεις και τις πλειάδες. Επιπλέον, έχει το πλεονέκτημα ότι δεν περιορίζεται από οποιοδήποτε λεξικό,

επειδή αφορά τις σχεσιακές βάσεις δεδομένων όπου τα σχήματα είναι προβλέψιμα και εξοικειωμένα σε έναν εμπειρογνώμονα, όπως για παράδειγμα ο διαχειριστής της βάσης (dba). Κατά συνέπεια ο μηχανισμός προτύπων που εισάγεται αργότερα σε αυτό το έγγραφο είναι ικανοποιητικός για το στόχο μας. Επιπλέον, οι ερωτήσεις *précis* είναι ερωτήσεις βασισμένες σε λέξεις - κλειδιά που μπορούν να οδηγήσουν σε σύνθετες ερωτήσεις SQL η των οποίων μορφή περιορίζεται μόνο από τη γραφική παράσταση του σχήματος της υπό χρήση βάσης δεδομένων (γράφος της βάσης). Τέτοιες εργασίες [21] χρησιμοποιούν ένα σύνολο προκαθορισμένων προτύπων ερώτησης, τα οποία δεν μπορούν να αξιωθούν πληρότητα, επειδή αυτό το σύνολο των ερωτήσεων είναι δύσκολο να υπολογίσει όλες τις πιθανές ερωτήσεις σε μια ορισμένη βάση δεδομένων. Επιπλέον, τα αποτελέσματα αυτών των εργασιών δίνουν προκαθορισμένες απαντήσεις, όπου μόνο οι τιμές στα πρότυπα σχέδια που χρησιμοποιούν αλλάζουν. Αυτό είναι σε αντίθεση με τις ερωτήσεις *précis*, οι οποίες κατασκευάζουν λογικά υποσύνολα κατ' απαίτηση και χρησιμοποιούν τα πρότυπα των προτάσεων που καθορίζονται στις ετικέτες της γραφικής παράστασης βάσεων δεδομένων (γράφος), παράγοντας κατά συνέπεια δυναμικές απαντήσεις. Αυτό το χαρακτηριστικό των ερωτήσεων *précis* επιτρέπει επίσης την χρήση πολλαπλών προτύπων. Σε αυτήν την εργασία, χτίσαμε επάνω στις προτεινόμενες ιδέες [11, 26] και υλοποιήσαμε την μετάφραση ενός λογικού υποσυνόλου βάσης δεδομένων που παράγεται από μια ερώτηση *précis* σε ένα αφηγηματικό κομμάτι των πληροφοριών.

3

Θεωρητικό υπόβαθρο

Ο σκοπός αυτού του κεφαλαίου είναι να παρασχεθούν οι βασικές πληροφορίες για τις ερωτήσεις *précis*. Καταρχάς, περιγράφουμε πώς μια βάση δεδομένων μπορεί να διαμορφωθεί ως γράφος, και περιγράφουμε ένα απλό παράδειγμα. Έπειτα, περιγράφουμε πρότυπα για επερωτήσεις *précis* και την αρχιτεκτονική του συστήματος μετάφρασής μας.

3.1 Εισαγωγικά

Εξετάζουμε τον γράφο $\mathbf{G}(\mathbf{V}, \mathbf{E})$ ενός σχήματος μιας βάσης δεδομένων ως έναν κατευθυνόμενο γράφο που αντιστοιχεί σε ένα σχήμα \mathbf{D} . Οι κόμβοι στο \mathbf{V} είναι:

- a) κόμβοι σχέσης, \mathbf{R} , ένας για κάθε σχέση στο σχήμα και
- b) κόμβοι πεδίων, \mathbf{A} , ένας για κάθε πεδίο κάθε σχέσης στο σχήμα.

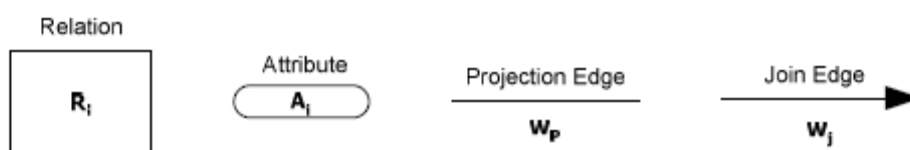
Οι ακμές στο \mathbf{E} είναι:

- a) ακμές προβολής (projection edges) Π , όπου καθεμία συνδέει έναν κόμβο πεδίου με τον αντίστοιχο κόμβο σχέσης που τον εμπεριέχει, η οποία αντιπροσωπεύει την πιθανή προβολή των πεδίων στην απάντηση και
- b) ακμές ένωσης, \mathbf{J} , από έναν κόμβο σχέσης σε έναν άλλο κόμβο σχέσης, αντιπροσωπεύοντας μια πιθανή ένωση μεταξύ αυτών των σχέσεων. Αυτές οι ενώσεις θα μπορούσαν να είναι ενώσεις που προκύπτουν φυσικά λόγω περιορισμών ξένων κλειδίων (foreign key constraints), αλλά θα μπορούσαν επίσης να είναι άλλες ενώσεις που έχουν νόημα σε κάποιον εξειδικευμένο χρήστη.

Οι ενώσεις είναι κατευθυνόμενες όπως εξηγούνται παρακάτω. Για απλότητα στην παρουσίαση, υποθέτουμε ότι

- a) τα αρχικά κλειδιά δεν είναι σύνθετα, κατά συνέπεια ένα πεδίο από μια σχέση ενώνει σε ένα μόνο πεδίο από μια άλλη σχέση, και
- b) αυτά τα πεδία έχουν το ίδιο όνομα. Το κοινό όνομα των πεδίων ένωσης επικολλάται στην αντίστοιχη ακμή των δύο εμπλεκόμενων σχέσεων.

Επομένως, μια γραφική παράσταση βάσεων δεδομένων ορίζεται τυπικά ως ένας κατευθυνόμενος γράφος $G(V,E)$, όπου: $V = R \cup A$, και $E = \Pi \cup J$. Η σημειογραφία για τη γραφική αναπαράστασή της δίνεται στο σχέδιο 3.1.1.

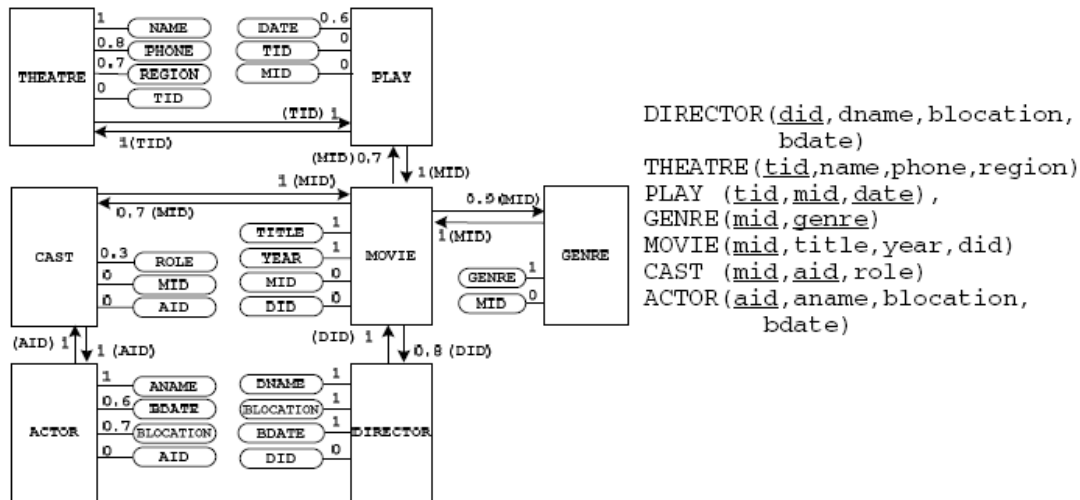


Σχήμα 3.1.1. Απεικόνιση των στοιχείων του γράφου

Ένα βάρος, w , ορίζεται σε κάθε ακμή του γράφου G . Αυτό είναι ένας πραγματικός αριθμός $[0, 1]$ που αντιπροσωπεύει την σπουδαιότητα της σύνδεσης μεταξύ των κόμβων. Ένα βάρος ίσο με 1 εκφράζει ισχυρή σχέση, με άλλα λόγια, εάν ένας κόμβος της ακμής εμφανίζεται σε μια απάντηση, τότε η άλλη άκρη θα πρέπει να ληφθεί υπόψη με αποτέλεσμα να εμφανιστεί και ο συνδεδεμένος κόμβος στην απάντηση. Εάν ένα βάρος είναι ίσο με 0, τότε η εμφάνιση ενός κόμβου στην απάντηση δεν προϋποθέτει την εμφάνιση της μετάφρασης και του συνδεδεμένου στην ακμή άλλου κόμβου. Δυο κόμβοι σχέσης θα μπορούσαν να συνδεθούν μέσω δυο διαφορετικών ακμών, στις δυο πιθανές κατευθύνσεις αλλά με διαφορετικά βάρη. Μια κατευθυνόμενη ακμή εκφράζει την εξάρτηση της τελικής σχέσης σε σχέση με την αρχική. Η αρχική σχέση εμπεριέχεται στην απάντηση και η τελική πιθανόν να περιληφθεί, αν ικανοποιεί την συνθήκη βάρους που έχουμε εφαρμόσει κάθε φορά. Για απλότητα στο συγκεκριμένο παράδειγμα, υποθέτουμε ότι υπάρχει το πολύ μια κατευθυνόμενη ακμή από έναν αρχικό κόμβο σε έναν τελικό κόμβο.

Ένα κατευθυνόμενο μονοπάτι μεταξύ δύο κόμβων σχέσης, περιλαμβάνοντας παρακείμενες ακμές, αντιπροσωπεύει έμμεσα την ένωση μεταξύ όλων αυτών των σχέσεων. Ομοίως, ένα κατευθυνόμενο μονοπάτι μεταξύ ενός κόμβου σχέσης και ενός κόμβου πεδίου, που περιλαμβάνει ένα σύνολο παρακείμενων ακμών ένωσης και μια ακμή προβολής αντιπροσωπεύει την έμμεση προβολή του πεδίου σε αυτήν την σχέση. Το βάρος ενός μονοπατιού είναι μια συνάρτηση του βάρους των προηγούμενων ακμών στο μονοπάτι. Αυτή

η συνάρτηση πρέπει να ικανοποιεί τον όρο ότι το βάρος μειώνεται όσο αυξάνεται το μονοπάτι, βασισμένο στην ανθρώπινη διαίσθηση και την λογική εξέλιξη της μετάφρασης.



Σχήμα 3.1.2. Παράδειγμα γράφου βάσης δεδομένων

Παράδειγμα.

Θεωρήστε μια βάση δεδομένων με ταινίες¹ που περιγράφεται από το σχήμα που παρουσιάζεται στο σχέδιο 3.1.2, όπου τα πρωτεύοντα κλειδιά υπογραμμίζονται. Παραδείγματος χάριν, παρατηρήστε τις δύο κατευθυνόμενες άκρες μεταξύ της σχέσης MOVIES και GENRE. Οι ταινίες και τα είδη τους συσχετίζονται αλλά κάποιος μπορεί να θεωρήσει ότι τα είδη εξαρτώνται περισσότερο από τις ταινίες παρά το ανάποδο. Με άλλα λόγια, μια απάντηση σχετικά με ένα είδος ταινίας θα πρέπει πάντα να περιέχει τις πληροφορίες για τις σχετικές ταινίες, ενώ μια απάντηση σχετικά με μια ταινία μπορεί να μην περιέχει απαραίτητως τις πληροφορίες για το είδος της. Για αυτόν τον λόγο, το βάρος της ακμής από το GENRE στην MOVIES τίθεται 1, ενώ το βάρος της ακμής από τον MOVIES στο GENRE τίθεται 0,9.

Χρησιμοποιώντας διαφορετικά βάρη στις ακμές του γράφου μας επιτρέπει να λαμβάνουμε διαφορετικές απαντήσεις στην ίδια ερώτηση. Τα βάρη μπορούν να παρασχεθούν με διάφορους τρόπους. Μπορούν να δοθούν από έναν χρήστη στο χρόνο ερώτησης χρησιμοποιώντας ένα κατάλληλο γραφικό περιβάλλον διασύνδεσης με τον χρήστη που επιτρέπει τη διαλογική εξερεύνηση του περιεχομένου μιας βάσης δεδομένων. Ένας χρήστης μπορεί να εξερευνήσει διαφορετικές περιοχές της βάσης δεδομένων ξεκινώντας, παραδείγματος χάριν, από εκείνη τη σχέση που σχετίζεται άμεσα με τις επιλογές της ερώτησής του και σταδιακά να επεκταθεί στα μέρη της βάσης δεδομένων που είναι πιο

¹ www.imdb.com

έμμεσα σχετιζόμενα με την ερώτησή του. Εναλλακτικά, σύνολα από βάρη που αντιστοιχούν σε διαφορετικές ερωτήσεις ή ομάδες χρηστών μπορούν να αποθηκευτούν στο σύστημα [20]. Παραδείγματος χάριν, αυτά τα διαφορετικά σύνολα θα ανταποκρίνονταν στις προτιμήσεις των κριτικών ή τους σινεφίλ ταινιών. Για το πρώτο, οι κριτικοί μπορεί να ενδιαφερθούν για λεπτομερείς σε βάθος απαντήσεις, όπου η χρήση ενός κατάλληλου συνόλου βαρών θα επέτρεπε σε αυτούς τους χρήστες για να ερευνήσουν ένα μεγάλο μέρος της βάσης δεδομένων γύρω από μια ενιαία ερώτηση *précis*. Από την άλλη, οι σινεφίλ προτιμούν συνήθως πιο σύντομες απαντήσεις. Σε αυτήν την περίπτωση, ένα διαφορετικό σύνολο βαρών θα επέτρεπε απαντήσεις που περιέχουν λιγότερες σχέσεις και αντικείμενα στην βάση. Τέλος, πολλά σύνολα βαρών που αντιστοιχούν σε διαφορετικούς χρήστες μπορούν να αποθηκευτούν στο σύστημα. Η χρήση διαφορετικών βαρών για κάποιο προφίλ χρήστη επιτρέπει εξατομικευμένες απαντήσεις. Παραδείγματος χάριν, ένας χρήστης μπορεί να ενδιαφερθεί για την περιοχή όπου ένα θέατρο βρίσκεται, ενώ ένας άλλος μπορεί να ενδιαφερθεί για το τηλέφωνο ενός θεάτρου. Εντούτοις, η προσέγγιση που παρουσιάζεται είναι γενική δεδομένου ότι δεν εξαρτάται από ένα συγκεκριμένο βάρος-πρότυπο.

3.2 Μοντέλο ερωτήσεων *précis*

Θεωρούμε μια βάση δεδομένων D κατάλληλα αρχικοποιημένη με ένα σύνολο βαρών και μια ερώτηση *précis*, η οποία αποτελείται από μια λέξη-κλειδί, δηλαδή $Q = \{k_1\}$. Ορίζουμε ως αρχική σχέση (*initial relation*) οποιαδήποτε σχέση της βάσης δεδομένων που περιέχει τουλάχιστον μια πλειάδα που περιέχει τη συγκεκριμένη λέξη κλειδί. Η πλειάδα στην οποία υπάρχει η λέξη-κλειδί ονομάζεται αρχική πλειάδα (*initial tuple*).

Ένα λογικό υποσύνολο D' της βάσης δεδομένων D ικανοποιεί τα εξής:

- Το σύνολο σχέσεων στο D' είναι ένα υποσύνολο των σχέσεων της αρχικής βάσης δεδομένων D .
- Για κάθε σχέση R_i' στο αποτέλεσμα D' , το σύνολο πεδίων του στο D' είναι ένα υποσύνολο των πεδίων του στο D .
- Για κάθε σχέση R_i' στο αποτέλεσμα D' , το σύνολο των πλειάδων του είναι ένα υποσύνολο του συνόλου των πλειάδων στην αρχική σχέση R_i' στο D (όταν προβάλλονται στο σύνολο πεδίων που είναι παρόντα στο αποτέλεσμα).

Το αποτέλεσμα της εφαρμογής της ερώτησης Q σε μια βάση δεδομένων D δεδομένου ενός συνόλου περιορισμών C είναι ένα λογικό υποσύνολο D' της βάσης δεδομένων D , έτσι ώστε το D' περιέχει αρχικές πλειάδες για το Q και οποιοδήποτε άλλη πλειάδα στο D που μπορεί να φτάσει μέσω συνδέσεων σχέσεων (διαμέσου των ξένων-κλειδιών) που αρχίζουν από κάποια

αρχική πλειάδα, υποκειμένες στους περιορισμούς C. Στους πιθανούς περιορισμούς στο C θα μπορούσε να περιλάβει το μέγιστο αριθμό ενώσεων, τον μέγιστο αριθμό πλειάδων στο D' και ούτω καθεξής. Η χρήση διαφορετικών περιορισμών επιτρέπει διαφορετικές απαντήσεις για την ίδια ερώτηση και για το ίδιο σύνολο βαρών στις ακμές του γράφου της βάσης δεδομένων. Όπως και στα βάρη, οι περιορισμοί μπορούν να διευκρινιστούν στο χρόνο ερώτησης, ή να προ-αποθηκευτούν στο σύστημα.

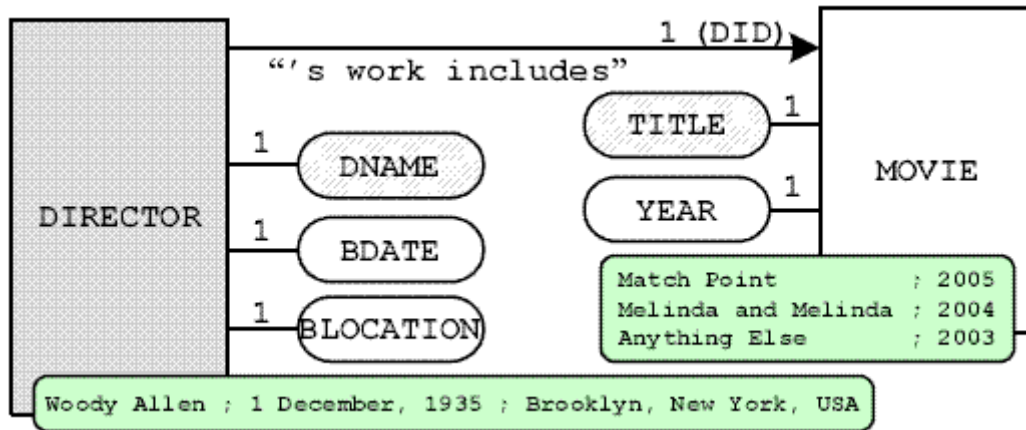
3.3 Λογική μετάφραση υποσυνόλων

Λαμβάνοντας υπόψη ένα λογικό υποσύνολο L για τα *precis queries* και ένα σύνολο περιορισμών που διευκρινίζουν τη μορφή της απάντησης, αυτή η φάση είναι υπεύθυνη για την ημιαυτόματη μετάφραση των πληροφοριών που αποθηκεύονται στο L σε μια φιλικότερη προς το χρήστη σύνθεση των αποτελεσμάτων. Στα πλαίσια αυτής της εργασίας, η παρουσίαση μιας απάντησης *query* ορίζεται ως μια σωστή δομημένη διαχείριση των μεμονωμένων αποτελεσμάτων, σύμφωνα με ορισμένους κανόνες και πρότυπα που προκαθορίζονται από το σχεδιαστή ή το διοικητή της βάσης δεδομένων. Δεν προσδοκούμε την κατασκευή ενός ευφυούς συστήματος. Μάλλον, δίνουμε μια φιλική προς το χρήστη απάντηση μέσω της σύνθεσης απλών προτάσεων.

Σε αυτό που ακολουθεί, περιγράφουμε αρχικά μερικές βασικές έννοιες που χρησιμοποιούνται στη μέθοδό μας και παρέχουμε έναν μηχανισμό προτύπων που διευκολύνει την κατασκευή των γενικών και εξατομικεύσιμων προτύπων ικανών για να παρουσιάσουν δυναμικά τα στοιχεία μιας ή περισσότερων σχέσεων. Κατόπιν, παρουσιάζουμε τον αλγόριθμο μεταφράσεών μας που χρησιμοποιεί αυτόν τον μηχανισμό προτύπων προκειμένου να συντεθεί ένα αφήγημα από πλειάδες που περιλαμβάνονται στο λογικό υποσύνολο.

3.3.1 Προκαταρκτικά

Προκειμένου να περιγραφεί η σημασιολογία μιας σχέσης R_i μαζί με τις ιδιότητές της στη φυσική γλώσσα, θεωρούμε ότι η R_i έχει μια εννοιολογική και φυσική έννοια. Η εννοιολογική έννοια μιας σχέσης εκφράζεται με ένα κατάλληλο σχόλιο που συνδέεται από με τη σχέση. Παραδείγματος χάριν, στο ΣΣΔΒΔ (σχεσιακό σύστημα διαχείρισης βάσεων δεδομένων) της Oracle μπορούμε να χρησιμοποιήσουμε τη δήλωση *COMMENT* για να προσθέσουμε ένα σχόλιο σε μια σχέση ή μια ιδιότητα. Για την απλότητα στην παρουσίαση, στη συνέχεια αυτού του τμήματος, αναφερόμαστε στην εννοιολογική έννοια μιας σχέσης με το όνομά της. Η φυσική έννοια μιας σχέσης αντιπροσωπεύεται από την αξία ενός από τα πεδία της που χαρακτηρίζει περισσότερο τις πλειάδες αυτής της σχέσης. Αυτή η ιδιότητα καλείται πρωτεύον πεδίο και απεικονίζεται με ένα σκιασμένο με στρογγυλεμένες γωνίες ορθογώνιο.



Σχήμα 3.3.1. Τμήμα του παραδείγματος Βάσης Δεδομένων

Παράδειγμα. Θεωρήστε ένα λογικό υποσύνολο βάσεων δεδομένων για μια ερώτηση χρηστών (π.χ., "Woody Allen") στη βάση δεδομένων *MOVIES* που περιλαμβάνει τις σχέσεις *DIRECTOR* (αρχική σχέση) και *MOVIE*, όπως διευκρινίζεται στο σχήμα 3.3.1. Αυτό το λογικό υποσύνολο περιλαμβάνει τις σχέσεις, τις ιδιότητες και τα βάρη που παρουσιάζονται στο σχήμα. Επίσης, περιέχει τέσσερις πλειάδες, μια στη σχέση *DIRECTOR* και τρεις στις σχέσεις *MOVIE*. Η σχέση *MOVIE* αντιπροσωπεύει εννοιολογικά "τις ταινίες" στον πραγματικό κόσμο κατά συνέπεια, ένα σχόλιο που περιέχει την έννοια *MOVIE* μπορεί να εκφράσει την εννοιολογική έννοιά του. Επιπλέον, "ο τίτλος" είναι το κύριο χαρακτηριστικό μιας "ταινίας" κατά συνέπεια, η σχέση *MOVIE* πρέπει να έχει το πεδίο *TITLE* ως πρωτεύον πεδίο, δεδομένου ότι ο "τίτλος" συλλαμβάνει τη φυσική έννοια μιας "ταινίας".

Πρωτεύοντα πεδία. Το πρωτεύον πεδίο h_R μιας σχέσης R ορίζεται ως το πεδίο του οποίου το όνομα αντιπροσωπεύει τη φυσική έννοια εκείνης της σχέσης. Εξ ορισμού, η ακμή προβολής που συνδέει ένα πρωτεύον πεδίο με την αντίστοιχη σχέση έχει ένα βάρος 1 και αυτή η ιδιότητα πρέπει πάντα να είναι παρούσα στο αποτέλεσμα ενός *precis query*. Ένας έμπειρος χρήστης κάνει την επιλογή των κύριων ιδιοτήτων στη φάση αρχικής κατασκευής της γραφικής παράστασης του σχήματος της βάσης δεδομένων.

Δε θεωρούμε ότι όλες οι σχέσεις πρέπει να έχουν ένα πρωτεύον πεδίο. Παραδείγματος χάριν, μια σχέση που χρησιμοποιείται μόνο για την αποθήκευση των σχέσεων *n-to-m* μεταξύ των διαφορετικών οντοτήτων (π.χ., η σχέση *CAST* στο σχήμα) δεν απαιτεί μια ιδιότητα τίτλων. Συνήθως, τέτοιες σχέσεις χρησιμοποιούνται μόνο για την κατασκευή πορειών που αντιπροσωπεύουν τις *query* απαντήσεις και δεν έχουν κανένα πεδίο στο λογικό υποσύνολο βάσεων δεδομένων (εκτός από εκείνα που χρησιμοποιούνται για τη σύνδεση των σχέσεων μέσω κλειδίων).

Ετικέτες. Κάθε ακμή προβολής e που ανήκει στο σύνολο Π που συνδέει μια ιδιότητα A_j^i με την περιλαμβανόμενη σχέση της R_i σχολιάζεται από μια ετικέτα που δηλώνει την έννοια, από την άποψη της φυσικής γλώσσας, της σχέσης μεταξύ αυτής της ιδιότητας και του πρωτεύοντος πεδίου της αντίστοιχης σχέσης. Εάν μια άκρη προβολής συνδέει έναν κόμβο σχέσης με τις ιδιότητες τίτλων της, κατόπιν η αντίστοιχη ετικέτα απεικονίζει τη σχέση αυτής της ιδιότητας με την εννοιολογική έννοια της σχέσης π.χ., "the *TITLE* of a *MOVIE*".

Κάθε ένωση άκρων e που ανήκουν στο σύνολο J μεταξύ δύο σχέσεων έχει μια ετικέτα που δηλώνει τη σχέση μεταξύ των ιδιοτήτων τίτλων των σχετικών σχέσεων π.χ., "το ΕΙΔΟΣ (*GENRE*) μιας ΤΑΙΝΙΑΣ (*TITLE*)". Η ετικέτα της ένωσης των άκρων που περιλαμβάνει μια σχέση χωρίς μια ιδιότητα τίτλων δηλώνει τη σχέση μεταξύ των προηγούμενων και επόμενων σχέσεων της.

Παράδειγμα. Όσον αφορά το σχέδιο του σχήματος 1, μια πιθανή ετικέτα που συνδέεται με την άκρη προβολής μεταξύ της σχέσης *MOVIE* και της ιδιότητάς της *YEAR* μπορεί να είναι: "the *YEAR* of a *MOVIE* (*TITLE*)". (Θυμίζουμε ότι το *TITLE* είναι το πρωτεύον πεδίο της σχέσης *MOVIE*.)

Περιορισμοί. Η μετάφραση των πληροφοριών που αποθηκεύονται σε μια ή περισσότερες σχέσεις μπορεί να πραγματοποιηθεί με διαφορετικά σχήματα. Μια σχέση μπορεί να περιγραφεί με τη χρήση είτε μόνο του πρωτεύοντος πεδίου της (τρόπος R^0) είτε όλων των πεδίων της (τρόπος R^*). Ομοίως, υπάρχουν περισσότεροι από ένας τρόποι να αντιπροσωπευθούν με τη μορφή κειμένου οι πληροφορίες στις πολλαπλές σχέσεις, συνδυάζοντας κάθε φορά τους διαφορετικούς τρόπους (R^0 ή R^*) των συμμετεχουσών σχέσεων. (Οι έννοιες αυτές αναλύονται με περισσότερη λεπτομέρεια αργότερα στην Ενότητα 5.1) Κατά συνέπεια, για τη φάση μετάφρασης, ένα σύνολο περιορισμών παρέχεται για να διευκολύνει την κατασκευή των απαντήσεων ακολουθώντας διαφορετικά μορφότυπα. Τέτοιοι περιορισμοί μπορούν να καθορίσουν:

- (α) το μήκος μιας φράσης σύμφωνα με τον αριθμό λέξεων ή γραμμών
- (β) ο μέγιστος αριθμός πλειάδων που πρέπει να υπολογιστεί προκειμένου να αποφασιστεί ποιος τρόπος είναι κατάλληλος και
- (γ) ο τρόπος μεταφράσεων π.χ., ολόκληρη η μετάφραση πρέπει να πραγματοποιηθεί στον τρόπο R^0 . Περισσότερες λεπτομέρειες σχετικά με τη λειτουργία και τη χρήση των περιορισμών δίνονται στην παρουσίαση του αλγορίθμου μεταφράσεών μας (επόμενο κεφάλαιο).

Προφίλ. Είναι δυνατό να υπάρχουν διαμορφωμένες διαφορετικές ρυθμίσεις (π.χ., διαφορετικές ιδιότητες τίτλων και ετικέτες στους κόμβους και τις άκρες) ανά προφίλ χρήστη ή ομάδας. Εντούτοις, αυτό δεν έχει επιπτώσεις στη δυνατότητα εφαρμογής της προσέγγισης που παρουσιάζεται, επειδή η επιλογή ενός προφίλ γίνεται προτού να πραγματοποιηθεί η μετάφραση. Κατά συνέπεια, στο υπόλοιπο αυτού του τμήματος, δεν συζητάμε αυτό το ζήτημα περαιτέρω.

3.3.2 Μηχανισμός προτύπων

Η σύνθεση των αποτελεσμάτων των ερωτήσεων ακολουθεί το σχήμα της βάσης δεδομένων και το συσχετισμό των σχέσεων μέσω των πρωτευόντων και ξένων κλειδιών. Επιπλέον, εμπλουτίζεται από τις αλφαριθμητικές εκφράσεις αποκαλούμενες ετικέτες προτύπων (template labels) που χαρτογραφούνται στις άκρες της γραφικής παράστασης του σχήματος βάσης δεδομένων.

Πρότυπα. Ορίζουμε ως ετικέτα l ενός κόμβου n την εννοιολογική έννοια του κόμβου που αποθηκεύεται ως σχόλιο στον συγκεκριμένο κόμβο, και την δείχνουμε ως $l(n)$. Παραδείγματος χάριν, η ετικέτα του κόμβου ιδιοτήτων *TITLE* είναι "title". Το όνομα ενός κόμβου πρέπει να καθοριστεί από το σχεδιαστή/διαχειριστή της βάσης δεδομένων.

Μια ετικέτα προτύπων, $label(u,z)$, ορίζεται σε κάθε άκρη $e(u,z)$ που ανήκει στο σύνολο E του γράφου $G(V,E)$ του σχήματος βάσης δεδομένων. Αυτή η ετικέτα χρησιμοποιείται για την ερμηνεία της σχέσης μεταξύ των τιμών των κόμβων u και του z σε ένα αφήγημα. Η ετικέτα προτύπων $label(u,z)$ μιας άκρης $e(u,z)$ τυπικά περιλαμβάνει τα ακόλουθα στοιχεία: (α) ένα μοναδικό προσδιοριστικό για την ετικέτα στον γράφο βάσης δεδομένων (β) το όνομα του κόμβου έναρξης ή πηγής, δηλ., $l(u)$, (γ) το όνομα του τελικού κόμβου ή κόμβου στόχου, δηλ., $l(z)$ και (δ) ένα πεπερασμένο σύνολο αλφαριθμητικών εκφράσεων. Μια απλή ετικέτα προτύπων μπορεί να έχει τη μορφή:

$$label(u, z) = expr_1 + l(u) + expr_2 + l(z) + expr_3 \quad (1)$$

όπου $expr_1$, $expr_2$, $expr_3$ είναι αλφαριθμητικές εκφράσεις και ο τελεστής "+" δρα ως τελεστής αλληλουχίας (σύνδεσης).

Προκειμένου να χρησιμοποιηθούν οι ετικέτες προτύπων ή για να καταχωρήσουμε νέες, χρησιμοποιούμε μια απλή γλώσσα για τα πρότυπα που υποστηρίζει τις μεταβλητές, τους βρόχους, τις λειτουργίες και τις μακροεντολές.

Σε ένα πρότυπο, όταν αναφερόμαστε στην εννοιολογική έννοια ενός κόμβου, χρησιμοποιούμε απλά το όνομά της. Όταν ένα στιγμιότυπο του κόμβου απαιτείται, χρησιμοποιούμε τον κόμβο ως μεταβλητή. Υπάρχουν δύο είδη μεταβλητών: μεταβλητές παραμέτρου και iterators βρόχων.

Μεταβλητές παραμέτρων. Οι μεταβλητές παραμέτρων είναι μαρκαρισμένες με το σύμβολο @ στην αρχή τους και αντικαθίστανται από τις τιμές σε χρόνο αρχικοποίησης. Σε πολλές περιπτώσεις, οι τιμές που επιστράφηκαν σε ένα αποτέλεσμα query από μια ορισμένη ιδιότητα θα μπορούσαν να είναι περισσότερες της μίας. Κατόπιν, χρησιμοποιούμε έναν κατάλογο παραμέτρων που δείχνονται όπως:

$$@ < parametername > []$$

Για τέτοιες λίστες, το μήκος τους θα πρέπει να παρέχεται σε χρόνο instantiation.

Παράδειγμα. Μια ετικέτα προτύπων για την άκρη προβολής $e(PHONE, THEATRE)$ θα μπορούσε να είναι:

```
label(PHONE, THEATRE) =  
  "The PHONE of the THEATRE  
  @THEATRE.NAME is @PHONE"
```

όπου το *PHONE* και το *THEATRE* αντιπροσωπεύουν την εννοιολογική έννοια των κόμβων (ιδιότητες και σχέση, αντίστοιχα) *PHONE* και *THEATRE*, π.χ., $l(PHONE) = \text{"phone"}$ και $l(THEATRE) = \text{"theatre"}$ αντίστοιχα. Επίσης, οι *@THEATRE.NAME* και *@PHONE* είναι μεταβλητές παραμέτρων με πιθανές τιμές "ALPHAVILLE" και "12345". Σε αυτήν την περίπτωση, μια έγκυρη ετικέτα για αυτήν την άκρη μπορεί να είναι η ακόλουθη:

```
"To τηλέφωνο του θεάτρου ALPHAVILLE είναι 12345"  
"The phone of the theatre is ALPHAVILLE 12345"
```

Μετρητές/Δείκτες βρόχων

Οι δείκτες βρόχων καθορίζονται στην δημιουργία περιορισμών των βρόχων, όπως θα αναλύσουμε αργότερα. Σε κάθε κύκλο του βρόχου, όλες οι εμφανίσεις του πεδίου αυτού που χρησιμοποιείται για τον ορισμό του βρόχου (δείκτης) στο σώμα του βρόχου αντικαθίστανται από την τρέχουσα αξία του (όπως παρόμοια μεταχειρίζονται οι δηλώσεις #DEFINE σε έναν προ- επεξεργαστή της C). Οι δείκτες οι οποίοι είναι μαρκαρισμένοι στο σώμα του βρόχου αντικαθίστανται, ακόμη και αν είναι μέρος μιας άλλης έκφρασης ή μεταβλητής. Ορίζουμε τέτοιες αντικαταστάσεις με το να εσωκλείσουμε το πεδίο αυτό (δείκτη) μεταξύ \$. Αυτή η λειτουργία επιτρέπει την παραπομπή όλων των τιμών μιας λίστας από παραμέτρους και διευκολύνει τη δημιουργία ενός αυθαίρετου αριθμού προσχηματισμένων εκφράσεων.

Συναρτήσεις :

Υλοποιούμε δυο συναρτήσεις :

arityOf (< list of parameters >)

που επιστρέφει το πλήθος από μια λίστα από παραμέτρους, κυρίως για τον ορισμό του άνω ορίου στους δείκτες των βρόχων.

cond(<condition>, <if true>, <if false>)

που περιέχει μια συνθήκη, όπου αντικαθίσταται το 2^ο όρισμα αν είναι αληθής, ή το 3^ο όρισμα αν είναι ψευδής η συνθήκη. Χρησιμοποιείται κυρίως πριν τις μακροεντολές (σε ακμές σύνδεσης), όπου η μακροεντολή μπορεί να επιστρέψει ένα ή πολλά αποτελέσματα.

Βρόχοι :

Οι βρόχοι γενικεύουν την ισχύ των προτύπων παρέχοντας στον σχεδιαστή την δυνατότητα να επεξεργαστεί πρότυπα με άγνωστο αριθμό μεταβλητών και με άγνωστο πλήθος αποτελεσμάτων για τις σχετικές παραμέτρους. Η γενική μορφή βρόχων είναι:

[< simple constraint >] {(loop body)}

Όπου <simple constraint> είναι μια απλή συνθήκη της μορφής :

<lower><operator><iterator><operator><upper>

Εξετάζουμε μόνο τη γραμμική αύξηση με βήμα αύξησης ίσο με 1. Ανώτερο και κατώτατο όριο (με αρχική τιμή 1) μπορεί να είναι συνάρτηση όπως η arityOf(), μεταβλητές ή σταθερές τιμές. Οι έγκυροι αριθμητικοί τελεστές σύγκρισης είναι +, <, >, =, <=, >= όλοι με τη συνηθισμένη σημασιολογία τους. Κατά τη διάρκεια των επαναλήψεων του βρόχου, το σώμα του βρόχου μεταφράζεται και συγχρόνως όλες οι μαρκαρισμένες εμφανίσεις του δείκτη του βρόχου αντικαθίστανται από την τρέχουσα αξία του, όπως περιγράφονται πριν.

Παράδειγμα : Θεωρήστε την ακόλουθη περίπτωση

```
[i ≤ arityOf(@MOVIE)] {@MOVIE[$i$]}
```

Το χαμηλότερο όριο έχει την αρχική τιμή 1, και το άνω όριο είναι ίσο με τον αριθμό των πεδίων της σχέσης MOVIES. Κατά συνέπεια, ο δείκτης i, παίρνει τιμές μεταξύ 1 και τον αριθμό των πεδίων της σχέσης αυτής. Όσον αφορά το βρόχο, περιέχει μια λίστα παραμέτρων που περιέχει τα πεδία της σχέσης MOVIES. Μετά την αξιολόγηση της συνάρτησης arityOf(MOVIE), ο βρόχος μπορεί να απεικονιστεί για παράδειγμα, ως

```
[i ≤ 2]{@MOVIE[$i$]}
```

Στην αντικατάσταση των μεταβλητών, στην συνέχεια, τα επόμενα αποτελέσματα λαμβάνονται :

```
MOVIE[$1$] = TITLE
```

```
MOVIE[$2$] = YEAR
```

Μακροεντολές :

Εισάγουμε τις μακροεντολές για να διευκολύνουμε τον ορισμό προτύπων και για να βελτιώσουμε την αναγνωσιμότητα τους. Οι μακροεντολές διευκολύνουν την επέκταση των πεδίων των σχέσεων και των μεταβλητών. Παραδείγματος χάριν, ένα σημαντικό πρόβλημα στον καθορισμό μιας γλώσσας για τα πρότυπα είναι η δυσκολία καθορισμού του αριθμού των πεδίων ή τις τιμές των πεδίων αυτών. Στο επίπεδο προτύπων, δεν είναι δυνατό να υπολογιστεί (α) ο αριθμός των πλειάδων που προβάλλονται στο λογικό υποσύνολο, και (β) τιμών των προς συζήτηση πλειάδων. Παραδείγματος χάριν, προκειμένου να ανακαλύψουμε :

(α) τα πεδία που προβάλλονται σε ένα λογικό υποσύνολο, χρειάζεται να δημιουργήσουμε μια σειρά από πεδία όπως ακολούθως :

```
DEFINE MOVIES_ATTR_LIST as
```

```
[i < arityOf(@MOVIE)]
```

```
{@MOVIE[$i$]}
```

```
[i = arityOf(MOVIE)]
```

```
{@MOVIE[$i$]}
```

(b) τους τίτλους μιας ταινίας που αντιστοιχούν σε μια συγκεκριμένη ερώτηση,

χρειάζεται να δημιουργήσουμε μια σειρά από τιμές όπως η ακόλουθη : DEFINE MOVIES_TUPLES_LIST as

```
[i < arityOf (@TITLE)]
```

```
{@TITLE[$i$]}
```

```
[i = arityOf(@TITLE)]
```

```
{@TITLE[$i$]}
```

Σημείωση : Παρατηρούμε την παρουσία δυο διαφορετικών βρόχων, προς αποφυγή ενός περιττού κόμματος “,” στο τέλος των προτάσεων.

Για το προαναφερθέν λογικό υποσύνολο του σχήματος 1, ο καθορισμός των πεδίων και μιας σειράς τιμών πεδίων εμφανίζεται στον πίνακα που ακολουθεί.

| | | | |
|------------------------|---|---|---|
| Σειρά από πεδία | <pre>DEFINE MOVIES_ATTR_LIST as [i < arityOf(@MOVIE)] {@MOVIE[\$i\$]} [i = arityOf(MOVIE)] {@MOVIE[\$i\$]}</pre> | <pre>DEFINE MOVIES_ATTR_LIST as [i < 2] {@MOVIE[\$1\$],} [i = 2] {@MOVIE[\$2\$]}</pre> | <pre>MOVIES_ATTR_LIST = { TITLE, YEAR }</pre> |
| Σειρά από τιμές πεδίων | <pre>MOVIES_TUPLES_LIST as [i < arityOf (@TITLE)] {@TITLE[\$i\$]} [i = arityOf(@TITLE)] {@TITLE[\$i\$]}</pre> | <pre>MOVIES_TUPLES_LIST as [i < 3] {@TITLE[\$1\$],} {@TITLE[\$2\$],} [i = 3] {@TITLE[\$i\$]}</pre> | <pre>MOVIES TITLES LIST = { “Match Point”, “Melinda and Melinda”, “Anything Else” }</pre> |

Ενδεικτικός πίνακας μακροεντολών.

3.3.3 Μετάφραση

Σε αυτήν την υπο- ενότητα, περιγράφουμε μια μέθοδο που αναλύει ένα λογικό υποσύνολο και συνθέτει ένα σύνολο αποτελεσμάτων ερωτήσεων σε μια αφηγηματική μορφή χρησιμοποιώντας το μηχανισμό προτύπων που παρουσιάστηκε προηγουμένως. Γενικά, ένα λογικό υποσύνολο μιας βάσης δεδομένων αποτελείται από μια αλληλουχία υπογράφων, καθένας από τους οποίους αντιστοιχεί σε μια διαφορετική ερμηνεία της ερώτησης που τίθεται στην βάση δεδομένων. Η μετάφραση πραγματοποιείται χωριστά για κάθε υπογράφο. Εάν κάποιος υπο- γράφος περιέχει μια αρχική σχέση, η μετάφραση αρχίζει από αυτήν την σχέση. Εάν ο υπο- γράφος περιέχει περισσότερες από μια αρχικές σχέσεις, τότε η μετάφραση αρχίζει από την αρχική σχέση που είναι η ρίζα στον αντίστοιχο αρχικό υπο- γράφο. Και στις δύο περιπτώσεις, η μετάφραση συνεχίζεται μετά από τη σύνδεση των σχέσεων μέσω των άκρων του υπογράφου. Κάθε πρόταση που δημιουργείται δίνει τις πληροφορίες που αποθηκεύονται σε μια σχέση όπως για παράδειγμα “Woody Allen was born in New York”, ή/και μεταφράζει τις σχέσεις μεταξύ των πλειάδων που αποθηκεύτηκαν στις σχέσεις που συνδέθηκαν μέσω ακμών σύνδεσης, όπως για παράδειγμα “Woody Allen has directed the movie “The Jade Scorpion”. Αυτές οι σχέσεις θα αναφέρονται ως “γείτονες”.

Η μετάφραση ολοκληρώνει όταν η διάσχιση του γράφου είναι πλήρης. Στο τέλος, όλες οι προτάσεις που παράγονται παρατίθενται. Σε αυτό που ακολουθεί, περιγράφουμε αρχικά τη μετάφραση των πληροφοριών που αποθηκεύονται σε μια ενιαία σχέση στηριζόμενοι στα αποτελέσματα του [26]. Κατόπιν, παρουσιάζουμε την κατασκευή των φράσεων που περιέχουν τις πληροφορίες που αποθηκεύονται σε πολλαπλές σχέσεις, με τη χρήση δυο γενικών ενοτήτων. Αυτές οι ενότητες συλλαμβάνουν πιθανές σχηματικές αντιπροσωπεύσεις που μπορούν να βρεθούν στην διάσχιση του γράφου. Τέλος, εισάγουμε έναν αλγόριθμο που χρησιμοποιεί τη μέθοδο που παρουσιάζεται προκειμένου να περιγραφεί τυπικά η μετάφραση ενός λογικού υποσυνόλου.

3.3.3.1 Μετάφραση μιας σχέσης

Η μετάφραση μιας ενιαίας σχέσης δημιουργείται από τις προτάσεις που κατασκευάζονται με τη χρήση των τιμών που είναι αποθηκευμένες σε αυτή. Για κάθε πλειάδα αυτής της σχέσης μια διαφορετική πρόταση δημιουργείται. Ο συσχετισμός όλων των προτάσεων δημιουργείται από την ετικέτα που ορίζεται στη σχέση από το χρήστη. Υπάρχουν δύο εναλλακτικές λύσεις για την αντιπροσώπευση του περιεχομένου μιας σχέσης:

(α) χρησιμοποιώντας μόνο το αρχικό πεδίο της σχέσης ή

(β) κατασκευάζοντας μια φράση που καθορίζεται από μακροεντολές που συνδυάζουν τις πληροφορίες που αποθηκεύονται στο πρωτεύον πεδίο μιας σχέσης, καθώς και στα υπόλοιπα πεδία της σχέσης που προβάλλονται στο λογικό υποσύνολο.

Η μετάφραση πληροφοριών που αποθηκεύεται σε μια σχέση R σχολιάζεται ως R° , όταν μόνο εξετάζεται η ιδιότητα τίτλων του R (περίπτωση α), και ως R^\bullet , όταν εξετάζονται οι πληροφορίες από όλες τις ιδιότητες του R (περίπτωση β).

Στην πρώτη περίπτωση, εάν ο τρόπος μεταφράσεων που επιλέγεται είναι R° , τότε η μεταφρασμένη πρόταση περιέχει μόνο το πρωτεύον πεδίο. Συνήθως, τέτοιες προτάσεις είναι υποδεέστερες προτάσεις, επομένως, δεν μπορούν να διαμορφώσουν χωριστές προτάσεις αλλά μπορούν να διαμορφώσουν μια πρόταση όταν ενώνονται με μια κύρια πρόταση - αυτό πραγματοποιείται στη μετάφραση πολλαπλάσιων σχέσεων (δείτε 1.3.2.) - . Στη δεύτερη περίπτωση, εάν ο τρόπος μεταφράσεων που επιλέγεται είναι R^\bullet , μια διαφορετική πρόταση ανά πεδίο πρέπει να δημιουργηθεί. Δεδομένου ότι το πρωτεύον πεδίο χαρακτηρίζει τις πληροφορίες που αποθηκεύονται σε μια σχέση, χρησιμοποιείται ως υποκείμενο και είναι το πρώτο μέρος σε κάθε μια από τις προτάσεις.

Έπειτα, οι ετικέτες των ακμών προβολής που συμμετέχουν στο λογικό υποσύνολο αξιολογούνται. Για πολλαπλά πεδία της ίδιας σχέσης, αναπόφευκτα, το ίδιο υποκείμενο (πρωτεύον πεδίο) πρέπει να επαναληφθεί αρκετές φορές. Για να αποφευχθεί αυτό, ένας έμπειρος χρήστης πρέπει να δημιουργήσει κατάλληλες εκφράσεις στις ετικέτες στις ακμές προβολής για να επιτρέψει στην κατασκευή σύνθετων σημαντικών προτάσεων. Αυτή η λειτουργία διευκολύνεται από έναν μηχανισμό εύρεσης κοινών εκφράσεων, ο οποίος βρίσκει τις κοινές εκφράσεις στις προτάσεις που αντιστοιχούν στην ετικέτα που συνδέεται με κάθε ακμή προβολής.

Παράδειγμα.

Για τη μετάφραση της σχέσης που απεικονίζεται στο παρακάτω σχήμα DIRECTORS σε μορφή R° , υποθέστε ότι οι ετικέτες των ακμών προβολής που συνδέουν το πεδίο DNAME, με τα πεδία BDATE και BLOCATION, τα οποία αποθηκεύουν τις πληροφορίες για την ημερομηνία γέννησης και τη τοποθεσία γέννησης ενός σκηνοθέτη, είναι οι ακόλουθες:

ετικέτα(h_R , BDATE) = @DNAME + "was born on" + @BDATE

ετικέτα(h_R , BLOCATION) = @DNAME + "was born in" + @BLOCATION

Όταν περιλαμβάνονται και τα δυο πεδία στην απάντηση, η εφαρμογή της εύρεσης κοινών εκφράσεων προσδιορίζει ότι το @DNAME και "was born" είναι κοινές εκφράσεις. Τέλος, η πρόταση που προέρχεται από την μετάφραση των παραπάνω μπορεί να είναι η ακόλουθη:

“@DNAME was born on @BDATE in @BLOCATION”

Τυπικά, η διαδικασία που χρησιμοποιείται για τη μετάφραση των πληροφοριών που αποθηκεύονται σε μια σχέση περιγράφεται από τον αλγόριθμο TR στο κεφάλαιο 5.

3.3.3.2 Μετάφραση πολλαπλών σχέσεων

Ένα λογικό υποσύνολο μιας βάσης δεδομένων περιέχει συνήθως περισσότερες από μια σχέσεις, και αποτελεί ένα γράφο. Δεδομένου ότι η διαδικασία μετάφρασης αναλύει αυτό το γράφο, μετά από το συσχετισμό των σχέσεων μέσω των πρωτευόντων και ξένων κλειδιών, δυο μορφές σύνδεσης μπορούν να προσδιοριστούν, όπως φαίνεται στο σχήμα 3.3.3.2.a . Ο πρώτος, M_U , αντιπροσωπεύει την απλή σύνδεση δύο διαδοχικών σχέσεων μέσω της σύνδεσης αναμετάξυ των (μοναδιαία ένωση).

Ο δεύτερος M_S , αντιπροσωπεύει τη σύνδεση μιας σχέσης πηγής με περισσότερες από μια σχέσεις στόχων (πολλαπλή ένωση).

Προκειμένου να περιγραφεί η μετάφραση πολλαπλών σχέσεων, περιγράφουμε αρχικά πώς αυτή η διαδικασία πραγματοποιείται σε καθεμία αυτών των μορφών σύνδεσης. Στην επόμενη υπο- ενότητα, παρουσιάζουμε έναν αλγόριθμο που περιγράφει τη μετάφραση ολόκληρου του γράφου.

Μοναδιαία ένωση (unary module)

Αρχικά εξετάζουμε την απλή περίπτωση δύο διαδοχικών (γειτονικών) σχέσεων R_i και R_j , που απεικονίζεται στο παρακάτω σχήμα ως μοναδιαία ένωση, M_U . Ουσιαστικά, εάν εξετάζουμε τη μοναδιαία ένωση αφαιρετικά, θα πρέπει να μεταφράσουμε τις πληροφορίες που υπάρχουν και στις σχέσεις R_i και R_j , και έπειτα να εκτιμήσουμε την ακμή που ενώνει τις δύο σχέσεις. Αυτή η μετάφραση αποτελείται από δύο μέρη: ένα μέρος που αντιπροσωπεύει τις πληροφορίες που αποθηκεύονται στη σχέση R_i και ένα μέρος που αντιστοιχεί στη συσχέτιση μεταξύ των δυο σχέσεων R_i και R_j . Κάθε ένα από αυτά τα μέρη μπορεί να περιλάβει περισσότερες από μια προτάσεις.

Το πρώτο μέρος κατασκευάζεται όπως συζητήθηκε στην προηγούμενη υπο- ενότητα. Εδώ εξηγούμε την κατασκευή των προτάσεων του δεύτερου μέρους. Προκειμένου να γίνει κατανοητή η σχέση μεταξύ των δύο εμπλεκόμενων σχέσεων, το πρωτεύον πεδίο της σχέσης R_i είναι το υποκείμενο της πρότασης που απασχολεί τη σύνδεση των σχέσεων. Το υπόλοιπο της πρότασης περιλαμβάνει κάποια έκφραση που και μια μεταβλητή- μακροεντολή που αντιπροσωπεύει τις πληροφορίες που αποθηκεύονται στη σχέση R_j . Η έκφραση αποθηκεύεται στην ετικέτα στην ακμή σύνδεσης των σχέσεων ενώνοντας το υποκείμενο της

σχέσης R_i με τις πληροφορίες της σχέσης R_j . Υπάρχουν δύο τρόποι να αντιπροσωπευθούν οι πληροφορίες που αποθηκεύονται στη σχέση R_j σε αυτό το πρότυπο:

- χρησιμοποιώντας μόνο το πρωτεύον πεδίο (R_j° μορφή) ή
- κατασκευάζοντας μια φράση χρησιμοποιώντας όλα τα πεδία (R_j^\bullet μορφή).

Στην προηγούμενη περίπτωση, προκειμένου να συμπληρωθεί η μετάφραση της σχέση- παιδί, μια πρόσθετη πρόταση αντιπροσωπεύοντας τις πληροφορίες που περιλαμβάνονται στα υπόλοιπα της σχέσης R_j απαιτείται. Αυτή η πρόταση κατασκευάζεται όπως περιγράφεται στην προηγούμενη υπο- ενότητα θεωρώντας ότι η σχέση- στόχος (παιδί) είναι μεταφρασμένη στον τρόπο R_j^\bullet . Η προαναφερθείσα ανάλυση μπορεί να συνοψιστεί τυπικά ως εξής :

Για τη μοναδιαία ενότητα M_U που περιλαμβάνει δύο σχέσεις R_i και R_j με κατεύθυνση από την πρώτη προς την τελευταία, η μετάφραση πραγματοποιείται σε δύο βήματα:

- (i) υπολογισμός της R_i^\bullet (όπως αναφέρεται στην προηγούμενη υπο- ενότητα)
- (ii) υπολογισμός είτε
 - ($R_i^\circ - R_j^\bullet$) \uplus R_j^\bullet ή
 - ($R_i^\circ - R_j^\circ$)

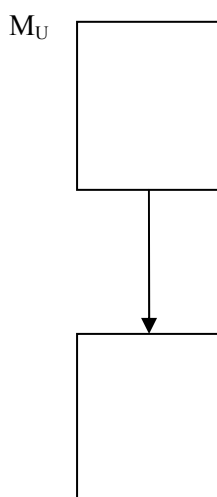
όπου $R_i^\circ - R_j^\bullet$ αποτελεί τη δημιουργία μιας φράσης που έχει ως μέρη τις πληροφορίες που αποσπώνται από τις R_i και R_j , η οποία αντιπροσωπεύεται από ένα κατάλληλο πρότυπο που υπάρχει στην σύνδεση των σχέσεων αυτών. Το σύμβολο \uplus αντιπροσωπεύει την σύνδεση διαφορετικών προτάσεων. Τελικά, οι προτάσεις που παράγονται από τα δύο αυτά βήματα συνδυάζονται για να παραγάγουν το τελικό αποτέλεσμα.

Συνεπώς, η μετάφραση μιας μοναδιαίας ένωσης παράγει αρχικά μια πρόταση που περιέχει τις πλήρεις πληροφορίες που αποθηκεύονται στη σχέση- πατέρα (βήμα (i)). Κατόπιν, υπάρχουν δύο εναλλακτικές λύσεις που ακολουθούν στο βήμα (ii). Μετά από τον πρώτο, ένα σύνολο φράσεων κατασκευάζεται αποτελούμενο από:

- (α) μια απλή φράση της μορφής "υποκείμενο - ρήμα - αντικείμενο" με την πιθανή προσθήκη μερικών σταθερών εκφράσεων για να συμπληρώσουν τη φράση, και
- (β) ένα κείμενο για την πλήρη μετάφραση της δεύτερης σχέσης της μορφής "υποκείμενο - ρήμα - διάφορες υπο- προτάσεις".

Ως εκ τούτου, παρατηρούμε στο βήμα (ii) ότι οι δύο εναλλακτικές λύσεις διατάζονται από την αυξανόμενη διάταξη της πολυπλοκότητας της κατασκευής. Η δεύτερη εναλλακτική λύση κατασκευάζει μια πιο συμπαγή, κατά συνέπεια πιο σύνθετη, και ελκυστική φράση. Εντούτοις, η παραγωγή μιας τέτοιας πρότασης δεν είναι πάντα εφικτή λόγω πιθανών περιορισμών που

επιβάλλονται σε κάποια περίπτωση, όπου και η πρώτη εναλλακτική λύση πρέπει να ακολουθηθεί. Γενικώς, η επιλογή μιας εκ των δυο γίνεται βάσει των περιορισμών που υπάρχουν.



Σχήμα 3.3.3.2.α Μοναδιαίας ένωση

Παράδειγμα.

Στο παράδειγμα του σχήματος 1, η σχέση DIRECTORS είναι συνδεδεμένη με τη σχέση MOVIE μέσω του κλειδιού DID. Αυτή η περίπτωση ταιριάζει με τη μοναδιαία ένωση. Το υποκείμενο της αντίστοιχης πρότασης για τη μετάφραση της ενιαίας σχέσης DIRECTORS θα είναι το πεδίο DNAME. Κατασκευάζουμε την πρόταση προτύπων που αντιστοιχεί στη σχέση DIRECTORS (βήμα (i)) όπως περιγράφεται στην προηγούμενη υπο- ενότητα:

ετικέτα(h_R , BDATE) = @DNAME + "was born on" + @BDATE

ετικέτα(h_R , BLOCATION) = @DNAME + "was born in" + @BLOCATION

Έπειτα, πρέπει να επιλέξουμε οποιαδήποτε από τις δύο εναλλακτικές λύσεις του βήματος (ii).

Θεωρούμε ότι οι περιορισμοί επιτρέπουν την επιλογή της δεύτερης από τις επιλογές : (R_i° - R_j^\bullet)

Θεωρούμε, ότι το αντίστοιχο πρότυπο για τη σχέση MOVIE είναι το ακόλουθο :

@TITLE + “(“ + @YEAR + “)”

Δεδομένου ότι η σχέση MOVIE μπορεί να περιέχει περισσότερες από μια πλειάδες, μια μακροεντολή απαιτείται για να επαναλάβει το ανωτέρω πρότυπο για όλες τις πλειάδες. Συνεχίζουμε με την πρόταση που συντίθεται από την ακμή που συνδέει τις σχέσεις DIRECTORS και MOVIE. Το πρότυπο αυτής της σχέσης μπορεί να αντιπροσωπευθεί από τον ακόλουθο τύπο:

ετικέτα (DIRECTOR, MOVIE) = @DNAME + expr1 + MOVIE_LIST

Η μακροεντολή MOVIE_LIST και οι εκφράσεις μπορούν να δηλωθούν ως εξής :

```
DEFINE MOVIE_LIST as
[i < arityOf(@TITLE)]
{@TITLE[$i$]+“ (“+@YEAR[$i$]+“),”}
[i = arityOf(@TITLE)]
f@TITLE[$i$]+“ (“+@YEAR[$i$]+“).”}
expr1 ← “’s work includes ”
```

Επομένως, μετά την αρχικοποίηση του προτύπου, και τις απαραίτητες συνδέσεις των προτάσεων, το αποτέλεσμα της ερώτησης précis για τον “Woody Allen”, για την σχέση DIRECTOR είναι:

*“Woody Allen was born on December 1, 1935 in
Brooklyn, New York, USA. Woody Allen’s work includes
Match Point (2005), Melinda and Melinda (2004), Anything Else (2003).”*

Εάν είχαμε εξετάσει την πρώτη εναλλακτική λύση του βήματος (ii), το αποτέλεσμα θα ήταν διαφορετικό δεδομένου ότι η μετάφραση των πληροφοριών που υπάρχουν στην σχέση MOVIE θα αντιπροσωπευόταν από περισσότερες από μια προτάσεις. Μια πιθανή ερμηνεία αυτού του παραδείγματος του σχήματος 3.3.1 είναι η ακόλουθη :

*“Woody Allen was born on December 1, 1935 in
Brooklyn, New York, USA. Woody Allen’s work includes
Match Point, Melinda and Melinda, Anything Else.
Match Point was released in 2005.
Melinda and Melinda was released in 2004.*

Anything Else was released in 2003."

Προφανώς, με τον μηχανισμό εύρεσης κοινών εκφράσεων που αναφέρεται στην προηγούμενη υπο- ενότητα, ένα πιο κομψό αποτέλεσμα θα παραγόταν. Εντούτοις, τα δύο αποτελέσματα έχουν μερικές κρίσιμες διαφορές. Ο πρώτος είναι πιο συμπαγής, δεν έχει οποιεσδήποτε επικαλύψεις, και μοιάζει με τη φυσική γλώσσα σε έναν μεγάλο βαθμό. Αφ' ετέρου, η δημιουργία της είναι πιο σύνθετη και σε μερικές περιπτώσεις ακόμα και απραγματοποίητη. Παραδείγματος χάριν, εάν η σχέση MOVIE στο λογικό υποσύνολο περιείχε περισσότερα πεδία και η δημιουργία περισσότερων από μιας πρότασης χρειάστηκε για να την περιγράψει, ένα τόσο κομψό αποτέλεσμα είναι δύσκολο να δημιουργηθεί. Το δεύτερο αποτέλεσμα κατασκευάζεται κατά τρόπο απλό και αποτελείται από σύνδεση διάφορων απλών φράσεων. Αυτό το είδος σύνθεσης μπορεί να περιγράψει τα πιο σύνθετα λογικά υποσύνολα.

Διαμεριστική ένωση (split module)

Μια άλλη περίπτωση προς εξέταση είναι ο συσχετισμός μεταξύ μιας ενιαίας σχέσης- πηγής και περισσότερων από μιας σχέσεων- παιδιών / στόχων, που απεικονίζονται στο παρακάτω σχήμα. Όπως πριν, η μετάφραση της ένωσης περιλαμβάνει δύο μέρη:

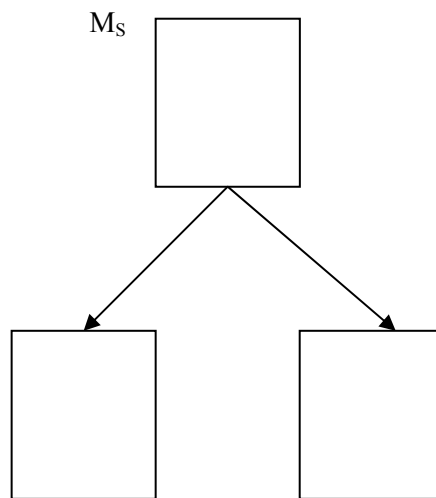
ένα που αντιστοιχεί στη σχέση- πατέρα / πηγής και μια που περιγράφει τη σύνδεση μεταξύ της σχέσης- πηγής και σχέσεων- στόχων. Το πρώτο μέρος κατασκευάζεται όπως περιγράφεται και στην προηγούμενη υπο- ενότητα (είναι κοινό και με τη μοναδιαία ένωση).

Το δεύτερο μέρος περιέχει ένα σύνολο προτάσεων περιλαμβάνοντας τα πρότυπα που υπάρχουν στις ακμές των σχέσεων που συνδέονται μεταξύ σχέσης- πηγής και σχέσεων- στόχων, μαζί με τις πληροφορίες που αποθηκεύονται στις σχέσεις- στόχων. Μια πρόταση που δημιουργείται χρησιμοποιώντας τις πληροφορίες που υπάρχουν στις ακμές σύνδεσης των διαφορετικών σχέσεων- στόχων έχει ως υποκείμενο το αρχικό πεδίο της σχέσης πηγής. Το υπόλοιπο της πρότασης περιλαμβάνει εκφράσεις και μια ή περισσότερες προτάσεις που αντιπροσωπεύουν τις πληροφορίες που αποθηκεύονται στις σχέσεις στόχων. Αν και είναι δυνατό να θεωρηθεί η περίπτωση ως δύο χωριστές μοναδιαίες ενώσεις, υπάρχει το πρόβλημα της αναπαραγωγής πληροφοριών από τη σχέση πηγής. Πολλαπλές μεταφράσεις τις σχέσης- πηγής είναι αναπόφευκτες. Προς αποφυγή αυτού του προβλήματος, οι πληροφορίες που των σχέσεων- στόχων παρουσιάζονται ως μια έκφραση που συνδυάζεται με μια λέξη ένωσης θ . Η προεπιλεγμένη λέξη στην προσέγγισή μας είναι η λέξη "και". Τα συστατικά αυτής της συνδυασμένης έκφρασης μπορούν να είναι είτε στον τρόπο R° είτε στον τρόπο μετάφρασης R^\bullet .

Κατά συνέπεια, τυπικά, για τη μετάφραση διαμεριστικής ένωσης περιλαμβάνοντας τις σχέσεις R_i και R_{j_1}, R_{j_2} , με την κατεύθυνση από την πρώτη προς τις επόμενες, μπορεί να γίνει στα εξής δυο βήματα :

- (iii) υπολογισμός της R_i^\bullet (όπως αναφέρεται στην προηγούμενη υπο- ενότητα)
- (iv) υπολογισμός είτε
 - $R_i^\circ - (R_{j_1}^\circ \theta R_{j_2}^\circ) \uplus R_{j_1}^\bullet \uplus R_{j_2}^\bullet$ ή
 - $R_i^\circ - (R_{j_1}^\bullet \theta R_{j_2}^\circ) \uplus R_{j_2}^\bullet$ ή
 - $R_i^\circ - (R_{j_1}^\bullet \theta R_{j_2}^\bullet)$

Παρατηρήστε ότι οι υποψήφιες επιλογές στο βήμα (ii) απαριθμούνται από την αυξανόμενη διάταξη πολυπλοκότητας της κατασκευής υπό την ίδια έννοια που περιγράφεται για τη μοναδιαία ένωση. Στο τέλος, τα αποτελέσματα και των δύο βημάτων συνδυάζονται για να παραγάγουν τελικό κείμενο. Η ανωτέρω ανάλυση ισχύει επίσης για μια διαμεριστική ένωση που αποτελείται από περισσότερες από δύο σχέσεις στόχων: R_{j_1}, \dots, R_{j_k} , με τους αντίστοιχους συνδυασμούς $R_{j_x}^\bullet$ ή $R_{j_x}^\circ$ μορφής.



Σχήμα 3.3.3.2.β Διαμεριστική ένωση

Παράδειγμα.

Εξετάστε το υποθετικό λογικό υποσύνολο που περιλαμβάνει τρεις σχέσεις, MOVIE (title, year), DIRECTOR(name, birth_location), και ACTOR (name, nationality) (τα πεδία title και name θεωρούνται ως πρωτεύοντα πεδία.) Επίσης, χάριν απλότητας στην παρουσίαση, τα κλειδιά παραλείπονται. Σε κάθε σχέση υπάρχει μια πλειάδα, και για κάθε σχέση το πρωτεύον πεδίο έχει την τιμή t_1^M , t_1^D , και t_1^A αντίστοιχα. Υποθέστε ότι το σχήμα αυτού του λογικού

υποσυνόλου είναι το ακόλουθο: $\text{DIRECTOR} \leftarrow \text{MOVIE} \rightarrow \text{ACTOR}$. Επομένως, αυτό το σενάριο μοιάζει με μια διαμεριστική ένωση, όπου R_i είναι η σχέση MOVIE ενώ R_{j_1} και R_{j_2} είναι οι άλλες δύο σχέσεις. Εξετάζοντας μερικές υποδειγματικές εκφράσεις στις ετικέτες των ακμών, καθεμία των προαναφερθεισών εναλλακτικών λύσεων για το βήμα (ii) μπορεί να παραγάγει τις προτάσεις που απεικονίζονται στον παρακάτω πίνακα. Παρατηρούμε ότι η πρώτη εναλλακτική λύση παράγει τις απλούστερες προτάσεις με κάποια επικάλυψη, ενώ η δεύτερη και κυρίως η τρίτη εναλλακτική πιο σύνθετα αλλά συνάμα και πιο κομψά αποτελέσματα μετάφρασης. Σε όλες τις περιπτώσεις, η λέξη προεπιλογής "και" χρησιμοποιείται ως λέξη ένωσης. Επίσης, στις τελευταίες δύο περιπτώσεις το πρότυπο που χρησιμοποιείται είναι ελαφρώς διαφορετικό όπως και αντίστοιχες εκφράσεις εμπλουτίστηκαν με τη λέξη "who".

| | | |
|-----------------|--|--|
| split module | $R_i^\circ - (R_{j_1}^\circ \theta R_{j_2}^\circ) \uplus R_{j_1}^\bullet \uplus R_{j_2}^\bullet$ | <i>The movie t_i^M involves the director t_i^D and the actor t_i^A.</i> <i>The director t_i^D was born in Italy. The actor t_i^A is Greek.</i> |
| | $R_i^\circ - (R_{j_1}^\bullet \theta R_{j_2}^\circ) \uplus R_{j_2}^\bullet$ | <i>The movie t_i^M involves the director t_i^D who was born in Italy and the actor t_i^A.</i> <i>The actor t_i^A is Greek.</i> |
| | $R_i^\circ - (R_{j_1}^\bullet \theta R_{j_2}^\bullet)$ | <i>The movie t_i^M involves the director t_i^D who was born in Italy and the actor t_i^A who is Greek.</i> |

3.3.4 Δυνατότητα εφαρμογής και περιορισμοί

Η παρουσιαζόμενη μέθοδος ισχύει σε οποιαδήποτε μορφή λογικού υποσυνόλου. Εντούτοις, η δυνατότητα εφαρμογής και η χρησιμότητα του μεταφραστή περιορίζονται από το μήκος των παραγόμενων αποτελεσμάτων. Ήδη τονίζεται ότι η πρόθεση αυτής της εργασίας δεν είναι η κατασκευή ενός ευφυούς συστήματος, καθώς αυτό δεν είναι εφικτό. Επομένως, η μέθοδός μας έχει περιορισμούς που σχετίζονται με το μήκος των τελικών αποτελεσμάτων. Ο στόχος είναι να διευκολυνθεί η παραγωγή απαντήσεων στις ερωτήσεις précis. Οι ερωτήσεις Précis σκοπεύουν να δώσουν χρήσιμες ιδέες στις ερωτήσεις των χρηστών. Μέσω της εμπειρίας μας στην αλληλεπίδραση με χρήστες, μια χρήσιμη απάντηση πρέπει να είναι σύντομη. Είναι πρακτικά άχρηστο να παραχθεί ένα αφηγηματικό κείμενο που περιέχει όλες τις γνωστές πληροφορίες για ένα σύνολο όρων. Είναι προτιμότερο να παραχθούν σύντομες και περιεκτικές απαντήσεις. Επομένως, ο περιορισμός στο μήκος του αποτελέσματος φαίνεται ως αναγκαίο. Αν και, υπάρχουν διάφοροι τρόποι να επιτύχει αυτό, αναφέρουμε τους πιο σημαντικούς.

- Να παρέχουμε στους χρήστες την ικανότητα να δηλώσουν περιορισμούς οι οποίοι ελέγχουν το μέγεθος της απάντησης, όπως εκείνους που έχουμε αναφέρει ήδη.
- Τη δυνατότητα για μετάφραση μόνο στον τρόπο R° , δηλαδή μόνο μετάφραση πληροφοριών που αφορούν το πρωτεύον πεδίο της προς μετάφραση σχέσης.
- Τη δυνατότητα να χωρίσουν τα αποτελέσματα σε περισσότερες από ένα κείμενο, όπως για παράδειγμα να κατασκευάζουν μια περιγραφή του W. Allen ως σκηνοθέτη και μια διαφορετική περιγραφή του W. Allen ως ηθοποιού.

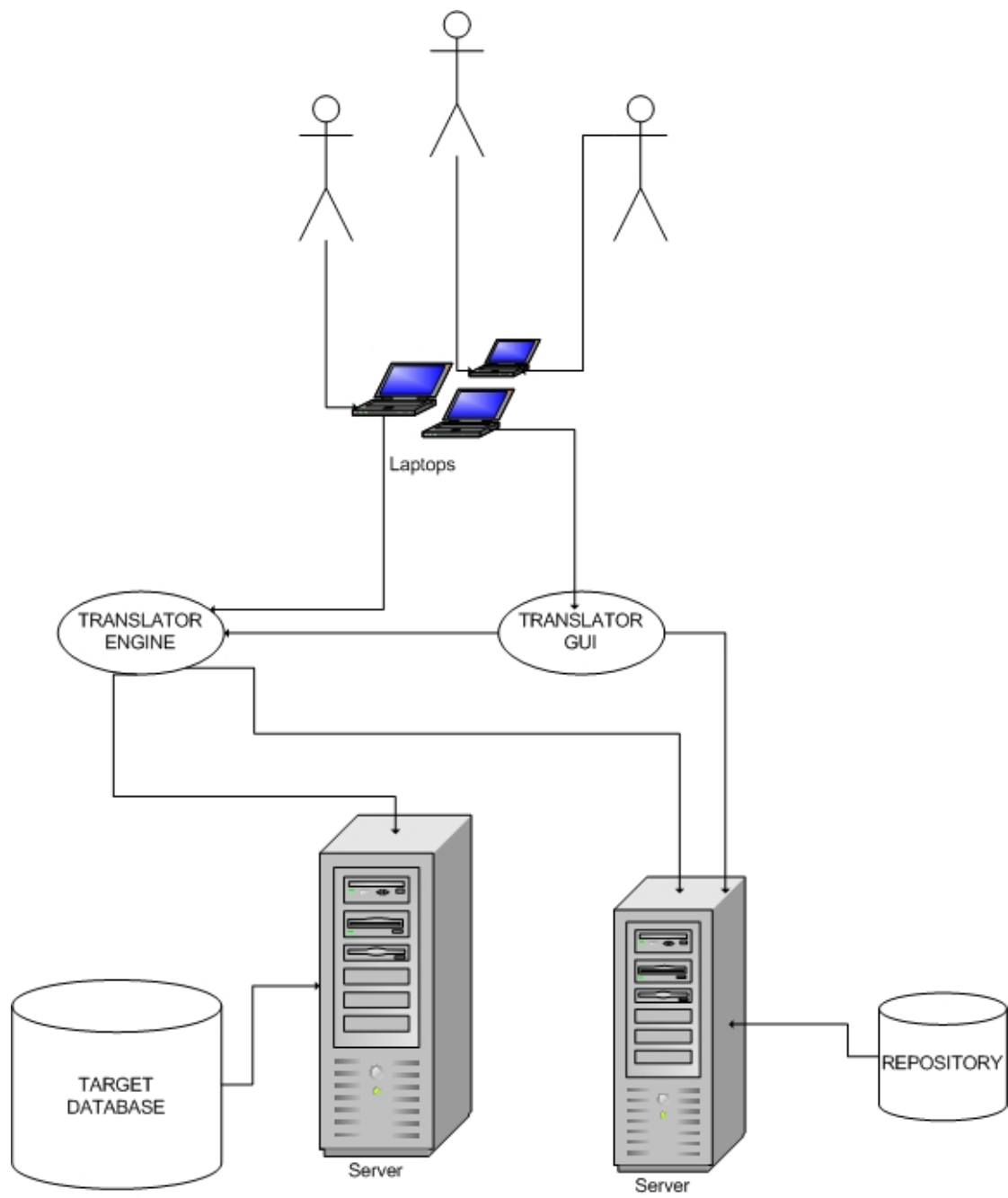
4

Σχεδίαση Συστήματος

Ακολουθεί η σχεδίαση του συστήματος που δημιουργήθηκε και το οποίο αποτελείται από τα εξής μέρη :

- i. Την μηχανή μετάφρασης, η οποία γράφτηκε σε γλώσσα C++
- ii. Την διεπαφή χρήστη (GUI), η οποία γράφτηκε σε γλώσσα Java και
- iii. Τις βάσεις δεδομένων, μια για αποθήκευση πληροφοριών σχετικών με την μετάφραση, και μια βάση από την οποία αντλούμε δεδομένα μετάφρασης.

Στο σχήμα 4.1 παρουσιάζεται η σχέση των συστατικών της εργασίας αυτής και πώς αυτά επικοινωνούν μεταξύ τους.



Σχήμα 4.1 Αρχιτεκτονική συστήματος

4.1 Αρχιτεκτονική

4.1.1 Μηχανή μετάφρασης

Η μηχανή μετάφρασης, η οποία γράφτηκε σε κώδικα C++, αναλαμβάνει τις λειτουργίες της μετατροπής της βάσης σε γράφο, τον εμπλουτισμό του γράφου αυτού με πληροφορίες που βρίσκονται στην βάση-repository (ετικέτες, μακροεντολές, καθώς και διάφορες πληροφορίες συστήματος σχετικά με την μετάφραση), και, τελικώς της διάσχισης του γράφου και την δημιουργία της τελικής απάντησης στον χρήστη σύμφωνα με τα δεδομένα του γράφου. Η ακριβής λειτουργία της μηχανής μετάφρασης εξηγείται αναλυτικά στο επόμενο κεφάλαιο.

Συνοπτικά, οι λειτουργίες που εκτελεί κάθε στοιχείο της μηχανής μετάφρασης, είναι :

- i. ***oraodbc.cpp*** : περιέχει την main μέθοδο, η οποία καλεί τις μεθόδους από τα υπόλοιπα μέρη και συγκεντρώνει όλες τις πληροφορίες για την μετάφραση. Η αυτόνομη λειτουργία που πραγματοποιεί είναι η υλοποίηση του αλγορίθμου *TRLS (Translation of a Logical Subset)*.
- ii. ***graph_impl.cpp*** : περιέχει όλες τις λειτουργίες για την δημιουργία του γράφου και τον εμπλουτισμό του με δεδομένα από την βάση-repository.
- iii. ***oracle_odbc.cpp*** : εδώ περιέχεται οτιδήποτε χρειάζεται για την επικοινωνία με τις βάσεις δεδομένων, όπως για παράδειγμα η αρχικοποίηση της σύνδεσης, η εκτέλεση μιας ερώτησης στην βάση και η αποσύνδεση από την βάση.
- iv. ***utils.cpp*** : διάφορες χρήσιμες λειτουργίες που χρησιμοποιούνται συχνά και από περισσότερα από ένα στοιχεία του κώδικα, οπότε και συγκεντρώθηκαν εδώ. Για παράδειγμα, η εύρεση του HA μιας σχέσης.
- v. ***parser.cpp*** : λειτουργίες που αφορούν γραμματική ανάλυση κειμένου βρίσκονται εδώ. Για παράδειγμα, οι λειτουργίες *resolve common expression* που βρίσκουν κοινές εκφράσεις/λέξεις σε ετικέτες, καθώς και ένας διαχωριστής κειμένου σε λέξεις, βρίσκονται εδώ
- vi. ***macro_parser.cpp*** : εδώ βρίσκεται οτιδήποτε σχετίζεται με τις μακροεντολές και την δημιουργία απαντήσεων από την ανάλυση των μακροεντολών αυτών και τον δεδομένων στην βάση
- vii. ***unary_module.cpp*** : η υλοποίηση της μετάφρασης μιας σχέσης, x και της σύνδεσής της με μια άλλη σχέση y , όπου η σχέση x έχει μια μόνο εξερχόμενη ακμή σύνδεσης προς την y και καμία άλλη. Εδώ μεταφράζονται οι ακμές προβολής της αρχικής σχέσης, και η ακμή σύνδεσης με την τελική σχέση.

- viii. ***split_module.cpp*** : εδώ υλοποιείται ο,τι και στο unary_module, με την διαφορά ότι η αρχική σχέση μπορεί να έχει παραπάνω από μια ακμή σύνδεσης με άλλες σχέσεις. Αυτό περιπλέκει την μετάφραση και για αυτόν τον λόγο δημιουργήθηκε αυτό το στοιχείο, μιας και περιέχει αρκετά τροποποιημένες τις λειτουργίες του unary_module.
- ix. ***globals.h*** : εκτός από τα header files για όλα τα στοιχεία κώδικα, υπάρχει και αυτό το ξεχωριστό header file, το οποίο, όπως υποδεικνύει και το όνομά του, περιέχει κάποιες παγκόσμιες μεταβλητές και σταθερές που είναι αναγκαίο να είναι ορατές από κάθε σημείο του κώδικα.

Στο σχήμα 4.1.1 φαίνεται η διάρθρωση του προγράμματος και η σύνδεση των επιμέρους στοιχείων μεταξύ τους.



Σχήμα 4.1.1 Αρχιτεκτονική μηχανής μετάφρασης

4.1.2 Γραφική διεπαφή χρήστη (GUI)

Η γραφική διεπαφή χρήστη, η οποία γράφτηκε σε γλώσσα προγραμματισμού Java, είναι ένα εργαλείο φιλικό προς έναν απλό χρήστη, με το οποίο κάποιος μπορεί να επισκοπήσει την βάση προς μετάφραση σε μορφή γράφου, να δημιουργήσει ετικέτες στους κόμβους του γράφου, να δημιουργήσει μακροεντολές, να διαμορφώσει κάποιες μεταβλητές συστήματος για την μετάφραση (όπως τι θα εμφανίζεται στην μετάφραση αν στην βάση κάποιο πεδίο είναι κενό και ούτω καθ' εξής), ή και να καλέσει την μηχανή μετάφρασης και να δει αποτελέσματα για τις επιλογές του. Επίσης, το εργαλείο αυτό είναι σχεδιασμένο να υποστηρίζει πολλά προφίλ χρηστών, ακόμα και οι χρήστες να βλέπουν τις επιλογές άλλων χρηστών, αλλά χωρίς να έχουν την δυνατότητα να τις μεταβάλλουν. Η ακριβής λειτουργία και λεπτομέρειες υλοποίησης του εργαλείου βρίσκονται στο επόμενο κεφάλαιο.

Συνοπτικά, οι λειτουργίες που εκτελεί κάθε κλάση της μηχανής μετάφρασης, είναι :

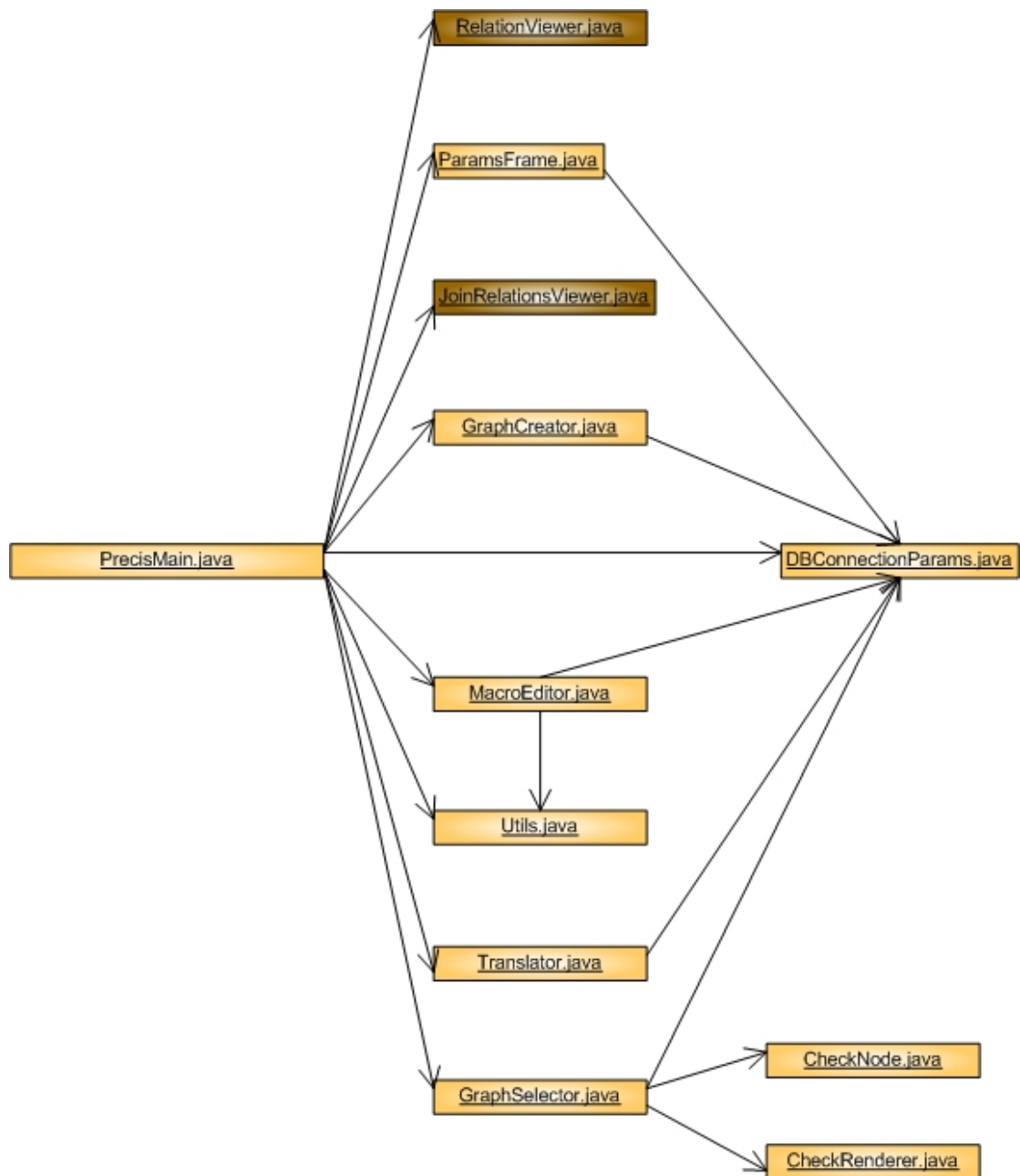
- i. **PrecisMain.java** : περιέχει την κύρια κλάση, η οποία παρουσιάζει την αρχική οθόνη για την εισαγωγή του χρήστη στο σύστημα, ζητώντας του την βάση για μετάφραση, την βάση-repository καθώς και όνομα χρήστη και κωδικό να συνδεθεί στο σύστημα. Εκτός αυτού, περιέχει όλα τα μενού για οποιαδήποτε άλλη οθόνη που χρειάζεται ο χρήστης για σύνταξη ετικετών, επισκόπηση γράφου, αλλαγή μεταβλητών συστήματος, καθώς και την επιλογή μετάφρασης. Μετά την εισαγωγή του χρήστη στο σύστημα, εμφανίζεται όλη η βάση σε μορφή γράφου στον χρήστη.
- ii. **RelationViewer.java** : είναι μια οθόνη που δείχνει τις σχέσεις και τα πεδία τους, και μπορεί κάποιος να δει τις ετικέτες στις ακμές προβολής κάποιας σχέσης, ή να θέσει κάποιο συγκεκριμένο πεδίο ως HA (πρωτεύον πεδίο). Εδώ μπορεί επίσης να δει μια συγκεκριμένη σχέση με τα πεδία της σε μορφή γράφου.
- iii. **RelationEditor.java** : εδώ πια μπορεί κάποιος να σβήσει από την βάση, να εισάγει ή να αλλάξει από την βάση μια ετικέτα που αντιστοιχεί σε κάποια ακμή προβολής. Μπορεί να αλλάξει το κείμενο ή το βάρος της ετικέτας. Επιπλέον, μπορεί να μεταβεί στην φόρμα δημιουργίας μακροεντολών για να κατασκευάσει μια μακροεντολή που θα περιέχει είτε μετα-δεδομένα της σχέσης αυτής, είτε μια μακροεντολή που θα περιέχει όλα τα πεδία της σχέσης με κείμενο της επιλογής του, σε περίπτωση που δεν επιθυμεί για κάθε πεδίο να φτιάχνει ξεχωριστή ετικέτα. Από αυτήν την φόρμα, τέλος, μπορεί να ενεργοποιήσει ή απενεργοποιήσει υπάρχουσες μακροεντολές, ανάλογα με τις επιθυμίες του (αν θέλει να μεταφράσει κείμενο βάσει των ετικετών ή των μακροεντολών).

- iv. **JoinRelationsViewer.java** : παρόμοια με την RelationViewer, αυτή η κλάση είναι μια οθόνη που δείχνει τις ετικέτες των ακμών σύνδεσης των σχέσεων. Με την ίδια λογική, μπορεί να δει τις εμπλεκόμενες σχέσεις στην ακμή αυτή, σε μορφή γράφου.
- v. **JoinRelationsEditor.java** : παρόμοια με την RelationEditor, μπορεί ο χρήστης να αλλάξει το κείμενο ή το βάρος της ετικέτας που περιέχει η ακμή σύνδεσης μεταξύ των δυο σχέσεων ή την επιλογή που επιτρέπει να γίνει override (να παραληφθεί από την μετάφραση) η σχέση- παιδί που ορίζεται από την ακμή αυτή. Επίσης μπορεί να κάνει αλλαγή γραμμής από εδώ, ή να προσθέσει μια συνάρτηση που αναγνωρίζεται από την μηχανή μετάφρασης και αφορά το πλήθος των εγγραφών στην τελική σχέση (αν είναι πληθυντικός ή ενικός, οπότε μπορεί να αλλάζουν κάποιες λέξεις), την $cond(arityOf()) > 1, <label\ if\ true>, <label\ if\ false>$).
- vi. **ParamsFrame.java** : αυτή είναι η φόρμα που περιέχει κάποιες παγκόσμιες μεταβλητές του συστήματος μετάφρασης, οι οποίες μπορεί να επιθυμεί ο χρήστης να αλλάξει (για παράδειγμα τι θα εμφανίζεται στην μετάφραση αν ένα πεδίο είναι κενό στην βάση).
- vii. **GraphCreator.java** : η υλοποίηση του γράφου σε γραφικά, για την εύκολη και αισθητικά κομψή επισκόπησή του από τον χρήστη
- viii. **MacroEditor.java** : μια περίπλοκη φόρμα που παρέχει στον χρήστη την δυνατότητα δημιουργίας μακροεντολών. Μέσα από αυτήν την φόρμα ο χρήστης μπορεί να κατασκευάσει μακροεντολές που αναφέρονται σε συνδέσεις σχέσεων, και οι οποίες είναι τρεις (κείμενο μαζί με είτε μόνο το πρωτεύον πεδίο, είτε μια επιλογή των πεδίων, είτε όλα), ή σε μετάφραση μιας σχέσης με τα πεδία της (δυο επιλογές, είτε μετα-δεδομένα της σχέσης, είτε μετάφραση πεδίων της). Η φόρμα αυτή παρέχει και αυτόματη δημιουργία μακροεντολών για την ευκολία του χρήστη.
- ix. **Utils.java** : είναι ένα Singleton αντικείμενο το οποίο χρησιμοποιούμε για να ελέγχουμε αν η φόρμα *MacroEditor* έχει ήδη ανοιχτεί από κάπου αλλού, ούτως ώστε να μην υπάρχουν δυο φόρμες *MacroEditor* ανοιχτές κάποια δεδομένη στιγμή.
- x. **Translator.java** : μια φόρμα που περιέχει τον κώδικα που είναι απαραίτητος να καλέσει την μηχανή μετάφρασης που έχει γραφτεί σε κώδικα C++, και να εμφανίσει τα αποτελέσματα στον χρήστη.
- xi. **GraphSelector.java** : μια φόρμα που δίνει στον χρήστη την δυνατότητα να παρουσιάσει όποιο μέρος της βάσης αυτός επιθυμεί, σε γράφο. Είτε είναι μόνο οι σχέσεις, είτε μια σχέση με τα πεδία της, είτε οποιαδήποτε μίξη αυτών, μπορεί να εμφανιστεί γραφικά στον χρήστη μέσα από αυτήν την φόρμα.

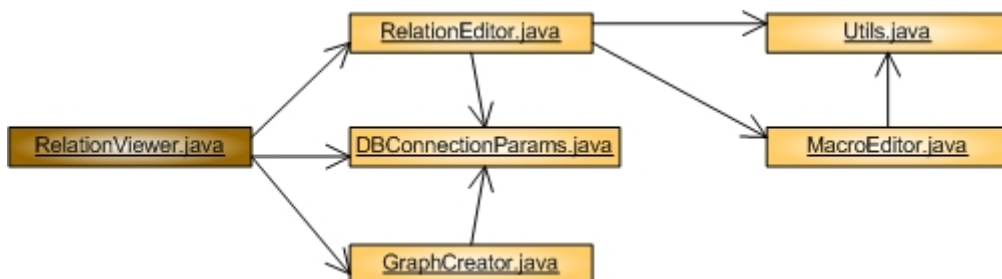
- xii. **CheckNode.java** : αυτή η κλάση χρησιμοποιείται από τον *GraphSelector* για να δημιουργήσουμε checkboxes στο δέντρο των σχέσεων και των πεδίων τους (η Java δεν παρέχει κάτι τέτοιο, έτσι χρειάστηκε να δημιουργηθεί.)
- xiii. **CheckRenderer.java** : όμοια, και αυτή η κλάση χρησιμοποιείται από τον *GraphSelector* για παρόμοιο λόγο, για το μέγεθος, το χρώμα και παραπλήσια χαρακτηριστικά των checkboxes του δέντρου, επειδή για τα checkboxes στο δέντρο δεν υπάρχει κάποιος *Renderer* στην Java.
- xiv. **DBConnectionParams.java** : αυτή η κλάση περιέχει όλες τις πληροφορίες για την σύνδεση με την βάση, τα ονόματα χρήστη και τους κωδικούς για τις δυο βάσεις στις οποίες συνδέεται, τις συνδέσεις και ο,τι έχει να κάνει με την διασύνδεση βάσης – GUI.

Σημείωση : οι εσωτερικές (inner) κλάσεις που χρησιμοποιήθηκαν στο γραφικό εργαλείο δεν αναφέρονται εδώ ούτε στα παρακάτω σχήματα για λόγους απλότητας.

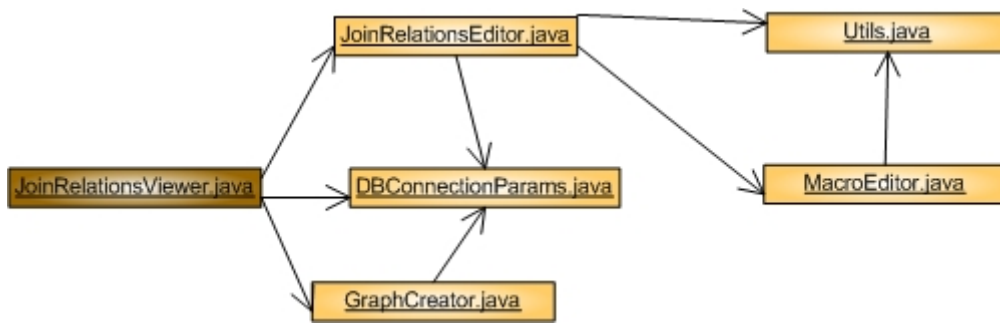
Στα σχήματα 4.1.2.1, 4.1.2.2 και 4.1.2.3 φαίνεται η διάρθρωση του προγράμματος και η σύνδεση των επιμέρους στοιχείων μεταξύ τους.



Σχήμα 4.1.2.1 Αρχιτεκτονική γραφικής διεπαφής χρήστη (GUI)



Σχήμα 4.1.2.1 Αρχιτεκτονική γραφικής διεπαφής χρήστη (GUI)



Σχήμα 4.1.2.3 Αρχιτεκτονική γραφικής διεπαφής χρήστη (GUI)

4.2 Περιγραφή Κλάσεων

Ακολουθούν οι περιγραφές του κώδικα για την μηχανή μετάφρασης και των κλάσεων για την γραφική διεπαφή χρήστη.

4.2.1 Μηχανή μετάφρασης

4.2.1.1 oraodbc.cpp

| Precis::oraodbc.cpp |
|--|
| void main(int argc, const char* argv[]); |

4.2.1.2 *graph_impl.cpp*

```
Precis::graph_impl.cpp
void InsertNodes(GRAPH <aNode, anEdge> * G1, aNode * arrNode,
                leda_list<leda_string> *ListArr, leda_string * type, int
                nodesColumn, int pKeysColumn);

void InsertProjNodeEdges(GRAPH <aNode, anEdge> * G1, aNode *
                        arrNode, anEdge * arrEdge,
                        leda_list<leda_string> *ListArr, leda_string *
                        type, int nodeSource, int nodeTarget, int
                        edgeName, int nodeComment);

void InsertJoinEdges(GRAPH <aNode, anEdge> * G1, aNode * arrNode,
                    anEdge * arrEdge, leda_list<leda_string> *ListArr,
                    leda_string * type, int nodeSource, int nodeTarget,
                    int join_p_key, int join_c_key);

void read_projection_info(GRAPH <aNode, anEdge> * G1,
                        leda_list<leda_string> *ListArr);

void read_join_info(GRAPH <aNode, anEdge> * G1,
                   leda_list<leda_string> *ListArr);
```

4.2.1.3 *oracle_odbc.cpp*

```
Precis::oracle_odbc.cpp
connVars connection( UCHAR *user, UCHAR *pass, UCHAR *dsn );

leda_list<leda_string> *ExecuteStatement( leda_string l_queryString,
                                         SQLHSTMT hstmt);
leda_list<leda_string> *ExecuteStatement( leda_string l_queryString,
                                         leda_string na,
                                         SQLHSTMT hstmt);

leda_list<leda_string> *ExecuteStatement( std::string l_queryString,
                                         SQLHSTMT hstmt);
leda_list<leda_string> *ExecuteStatement(std::string l_queryString,
                                         leda_string na,
                                         SQLHSTMT hstmt);

void disconnect( connVars * l_connVars );
```

4.2.1.4 *utils.cpp*

```
Precis::utils.cpp
leda_string stringToLeda(std::string inString);

std::string ledaToString(leda_string inString);

leda_string retrieveSingleTupleFromRepository(leda_string query, int
                                             column, int row);

leda_string listContent(aNode * aNodeP, int listType);

leda_string getHAVariable(leda_node currNode);
```

4.2.1.5 *parser.cpp*

```
Precis::parser.cpp  
leda_list<leda_string> tokenizer( leda_list<leda_string> * delims,  
                                leda_string sentence, int maxPosition);  
  
leda_list<leda_string> resolveCommonExpressions(  
    leda_list<leda_string> * sentenceList,  
    leda_string headAttr, int resolveMode);  
leda_list<leda_string> resolveCommonExpressions(  
    leda_list<leda_string> * sentenceList,  
    leda_string headAttr, leda_string conj_word,  
    int resolveMode);
```

4.2.1.6 *macro_parser.cpp*

```
Precis::macro_parser.cpp  
  
void generate_answer();  
  
void clear_globals(char const* first, char const* last);  
  
void vectorize(char const* first, char const* last);  
  
void vectorize_temp(char const* first, char const* last);  
  
void parse_existent_attrs(char const* first, char const* last);  
  
void set_def_bound(char const* first, char const* last);  
  
void set_att_counter(char const* first, char const* last);  
  
void set_function(char const* first, char const* last);  
  
void set_op(char const* first, char const* last);  
  
void set_iterator(char const* first, char const* last);  
  
void parse_loop_content(char const* first, char const* last);  
  
void iterator_atts(char const* first, char const* last);  
  
void header(char const* first, char const* last);  
  
std::string parse_macro(string macro, string * where_clause, connVars *  
    l_connVars_hr,  
    connVars * l_connVars_info, aNode * aNodeP,  
    int typeOfTranslation);  
std::string parse_macro(string macro, string * where_clause, connVars *  
    l_connVars_hr,  
    connVars * l_connVars_info, aNode * aNodeP,  
    int typeOfTranslation, int specificTuple);  
  
std::string parse_macro(string macro, connVars * l_connVars_hr,  
    connVars * l_connVars_info,  
    aNode * aNodeP, int typeOfTranslation);  
  
leda_list<leda_string> findAttributes(string macro);
```

4.2.1.7 unary_module.cpp

```
Precis::unary_module.cpp  
leda_string createUnaryJoinSentence(GRAPH <aNode, anEdge> * G1,  
    leda_edge * e, connVars *  
    l_connVars_hr,  
    connVars * l_connVars_info,  
    leda_string * ha_name,  
    leda_string * where_clause);  
  
leda_string unaryTranslate(GRAPH <aNode, anEdge> * G1,  
    leda_node * currNode,  
    leda_string * relationName,  
    connVars * l_connVars_hr,  
    connVars * l_connVars_info, int * tupID,  
    int mode,  
    leda_string *where_clause);
```

4.2.1.8 split_module.cpp

```
Precis::split_module.cpp  
leda_string splitTranslate(GRAPH <aNode, anEdge> * G1,  
    leda_node * currNode,  
    leda_string * relationName,  
    connVars * l_connVars_hr,  
    connVars * l_connVars_info,  
    int * tupID, leda_string *where_clause);  
  
leda_string createSplitJoinSentence(GRAPH <aNode, anEdge> * G1,  
    leda_list<leda_edge> * e,  
    connVars * l_connVars_hr,  
    connVars * l_connVars_info,  
    leda_string * ha_name, leda_string *  
    where_clause);  
  
void splitAppendGlobalWheres( GRAPH <aNode, anEdge> * G1,  
    leda_edge * e,  
    leda_string * where_clause,  
    leda_list<myleda_string_list> * list,  
    leda_list<myleda_node_list> * nodeList,  
    int globalWheresMixedListIterator);
```

4.2.2 Γραφική διεπαφή χρήστη

4.2.2.1 *PrecisMain.java*

| <i>precisgui.PrecisMain</i> |
|---|
| <i>Attributes</i> |
| <pre>public javax.swing.JMenuItem Exit; public javax.swing.JPanel ReposConnPanel; public javax.swing.JButton connectionButton; public javax.swing.JPanel connectionPanel; public javax.swing.JPanel dbConnectionPanel; public javax.swing.JMenuItem editJoinEdge; public javax.swing.JMenuItem editParamsMenu; public javax.swing.JMenuItem editProjEdge; public javax.swing.JSeparator fifthSeparator; public javax.swing.JSeparator firstSeparator; public javax.swing.JSeparator fourthSeparator; public javax.swing.JLabel hostLabel; public javax.swing.JMenuItem macroEditorMenu; public javax.swing.JMenu myMenu; public javax.swing.JMenuBar myMenuBar; public javax.swing.JMenuItem newConnection; public javax.swing.JPasswordField passField; public javax.swing.JLabel passLabel; public javax.swing.JPanel profilePanel; public javax.swing.JLabel profileidNameLabel; public javax.swing.JPasswordField profileidPassword; public javax.swing.JLabel profileidPasswordLabel; public javax.swing.JTextField profileidUserNameField; public javax.swing.JLabel reposHostLabel; public javax.swing.JPasswordField reposPassField; public javax.swing.JLabel reposPassLabel; public javax.swing.JTextField reposUrlField; public javax.swing.JTextField reposUserNameField; public javax.swing.JLabel reposUserNameLabel; public javax.swing.JSeparator secondSeparator; public javax.swing.JMenuItem specificGraph; public javax.swing.JSeparator thirdSeparator; public javax.swing.JMenu translateMenu; public javax.swing.JMenuItem translateMenuItem; public javax.swing.JTextField urlField; public javax.swing.JTextField userNameField; public javax.swing.JLabel userNameLabel; private DBConnectionParams dbConnParams = null; private Connection connection = null; private Connection infoConnection = null; private RelationViewer relViewer = null; private JoinRelationsViewer joinRelViewer = null; private MacroEditor macroEditor = null; private ParamsFrame params = null; private GraphCreator gc = null; private GraphSelector gSelector = null; private GLPanel glPanel = null; private Utils utils = null;</pre> |

precisgui.PrecisMain

Operations

```
public PrecisMain()
private void initComponents()
private void translateMenuItemActionPerformed(java.awt.event.ActionEvent evt)
private void specificGraphActionPerformed(java.awt.event.ActionEvent evt)
private void editParamsMenuActionPerformed(java.awt.event.ActionEvent evt)
private void macroEditorMenuActionPerformed(java.awt.event.ActionEvent evt)
private void newConnectionActionPerformed(java.awt.event.ActionEvent evt)
private void editJoinEdgeActionPerformed(java.awt.event.ActionEvent evt)
private void editProjEdgeAction(java.awt.event.ActionEvent evt)
private void connectActionPerformed(java.awt.event.ActionEvent evt)
private void exitAction(java.awt.event.ActionEvent evt)
public static void main(String args[])
```

4.2.2.2 *RelationViewer.java*

| <i>precisgui.RelationViewer</i> |
|--|
| <i>Attributes</i> |
| <pre>public javax.swing.JButton editTemplateButton; public javax.swing.JButton graphButton; public javax.swing.JPanel labelPanel; public javax.swing.JScrollPane labelScrollPane; public javax.swing.JPanel optionPanel; public javax.swing.JTextArea projEdgeArea; public javax.swing.JPanel relPanel; public javax.swing.JButton setHAButton; public javax.swing.ButtonGroup translationTypeGroup; // End of variables declaration private JTree relationsTree = null; private TreePath currentPath = null; private String currentPathComp = null; private Connection connection = null; private Connection infoConnection = null; private HashMap hAttributes = new HashMap(); DBConnectionParams dbConnParams = null; RelationEditor relEditor = null; public TreeSelectionListener treeSelListener = new TreeSelectionListener(){..}</pre> |
| <i>Operations</i> |
| <pre>public RelationViewer(DBConnectionParams dbConnParams, Vector nodeVector) private void initComponents() private void editTemplateButtonActionPerformed(java.awt.event.ActionEvent evt) private void graphButtonActionPerformed(java.awt.event.ActionEvent evt) private void setHAButtonActionPerformed(java.awt.event.ActionEvent evt) private void showArea() private void showGraph() private JTree createSourceTree(Vector nodeVector) public JTree getRelationsTree() public void setRelationsTree(JTree relationsTree) public HashMap getHAttributes() public void setHAttributes(HashMap hAttributes)</pre> |

4.2.2.3 *RelationEditor.java*

| <i>precisgui.RelationEditor</i> |
|--|
| <i>Attributes</i> |
| <pre>private javax.swing.JButton addHaButton; private javax.swing.JButton addPlusButton; private javax.swing.JButton addVarButton; private javax.swing.JPanel buttonPanel; private javax.swing.JButton clearAllButton; private javax.swing.JButton delTemplateButton; private javax.swing.JPanel editPanel; private javax.swing.JLabel haLabel; private javax.swing.JTextArea labelArea; private javax.swing.JScrollPane labelAreaScrollPane; private javax.swing.JButton macroButton; private javax.swing.JLabel macroLabel; private javax.swing.JPanel optionsPanel; private javax.swing.JPanel relPanel; private javax.swing.JButton saveTemplateButton; private javax.swing.JTextField weighthField; private javax.swing.JPanel weighthPanel; DBCConnectionParams dbConnParams = null; private Connection connection = null; private Connection infoConnection = null; private HashMap hAttributes = null; private boolean tupleExists = false; private JMenuItem[] items = null; private MacroEditor macroEditor = null; public TreeSelectionListener treeSelListener = new TreeSelectionListener(){..} private ActionListener macroSelectionListener = new java.awt.event.ActionListener(){..}</pre> |
| <i>Operations</i> |
| <pre>public RelationEditor(RelationViewer previousFrame, DBCConnectionParams dbConnParams, JTree relationsTree) private void initComponents() private void addPlusButtonActionPerformed(java.awt.event.ActionEvent evt) private void addVarButtonActionPerformed(java.awt.event.ActionEvent evt) private void addHaButtonActionPerformed(java.awt.event.ActionEvent evt) private void clearAllButtonActionPerformed(java.awt.event.ActionEvent evt) private void macroEditorMenuActionPerformed(java.awt.event.ActionEvent evt) private void macroSelectorMenuActionPerformed(String listName) private void resetMacroItemActionPerformed(boolean showResult) private void macroButtonActionPerformed(java.awt.event.ActionEvent evt) private void delTemplateButtonActionPerformed(java.awt.event.ActionEvent evt) private void saveTemplateButtonActionPerformed(java.awt.event.ActionEvent evt)</pre> |

4.2.2.4 *JoinRelationsViewer.java*

| <i>precisgui.JoinRelationsViewer</i> |
|---|
| <i>Attributes</i> |
| <pre>public javax.swing.JButton editTemplateButton; public javax.swing.JButton graphButton; public javax.swing.JTextArea joinEdgeArea; public javax.swing.JPanel labelPanel; public javax.swing.JScrollPane labelScrollPane; public javax.swing.JPanel optionPanel; public javax.swing.JPanel sourceRelPanel; public javax.swing.JPanel targetRelPanel; protected JTree sourceTree = null; protected JTree targetTree = null; private Connection connection = null; private Connection infoConnection = null; DBConnectionParams dbConnParams = null; JoinRelationsEditor joinRelEditor = null; public TreeSelectionListener treeSelListener = new TreeSelectionListener(..);</pre> |
| <i>Operations</i> |
| <pre>public JoinRelationsViewer(DBConnectionParams dbConnParams, Vector nodeVector) private void initComponents() private void editTemplateButtonActionPerformed(java.awt.event.ActionEvent evt) private void graphButtonActionPerformed(java.awt.event.ActionEvent evt) private void showArea() private void showGraph() private JTree createTree(Vector nodeVector) private void enablePanel(JPanel panel) private void disablePanel(JPanel panel) private JTree createTree(Vector nodeVector)</pre> |

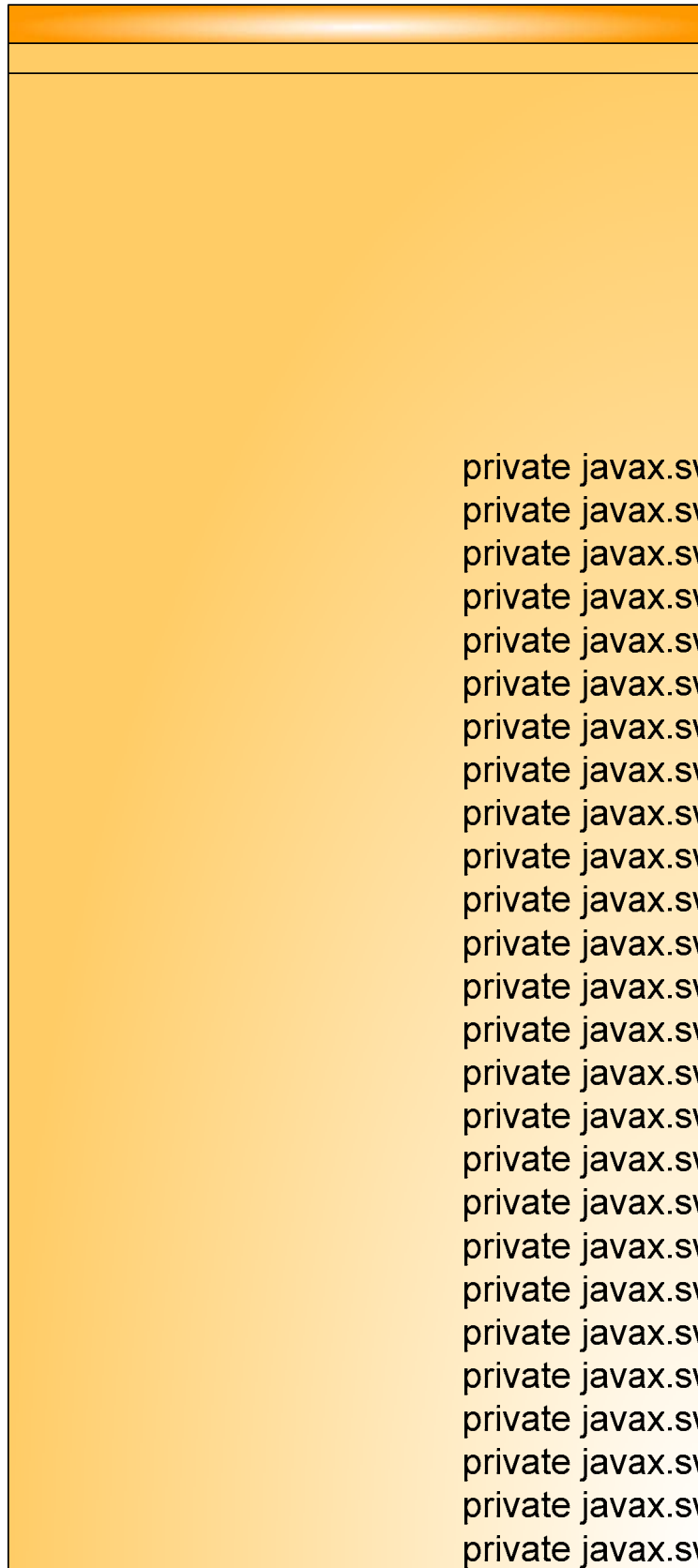
4.2.2.5 *JoinRelationsEditor.java*

| <i>precisgui.JoinRelationsEditor</i> |
|---|
| <i>Attributes</i> |
| <pre>public javax.swing.JPanel addButtonPanel; public javax.swing.JPanel buttonPanel; public javax.swing.JButton condButton; public javax.swing.JTextField conjunctiveWordField; public javax.swing.JLabel conjunctiveWordLabel; public javax.swing.JButton delTemplateButton; public javax.swing.JButton eolButton; public javax.swing.JTextArea joinEdgeArea; public javax.swing.JPanel labelPanel; public javax.swing.JScrollPane labelScrollPane; public javax.swing.JButton macroEditorButton; public javax.swing.JPanel modulePanel; public javax.swing.JPanel optionPanel; public javax.swing.JCheckBox overrideBox; public javax.swing.JButton saveTemplateButton; public javax.swing.JPanel sourceRelPanel; public javax.swing.JPanel targetPanel; public javax.swing.JPanel targetRelPanel; public javax.swing.JComboBox targetSelectorComboBox; public javax.swing.JTextField weighthField; public javax.swing.JPanel weighthPanel; protected JTree sourceTree = null; protected JTree targetTree = null; private Connection connection = null; private Connection infoConnection = null; private boolean tupleExists = false; DBConnectionParams dbConnParams = null; private MacroEditor macroEditor = null; private Params params = null; public TreeSelectionListener treeSelListener = new TreeSelectionListener(..);</pre> |
| <i>Operations</i> |
| <pre>public JoinRelationsEditor(JoinRelationsViewer previousFrame, DBConnectionParams dbConnParams, JTree srcRelationsTree, JTree trgtRelationsTree) private void initComponents() private void macroEditorButtonActionPerformed(java.awt.event.ActionEvent evt) private void eolButtonActionPerformed(java.awt.event.ActionEvent evt) private void condButtonActionPerformed(java.awt.event.ActionEvent evt) private void delTemplateButtonActionPerformed(java.awt.event.ActionEvent evt) private void saveTemplateButtonActionPerformed(java.awt.event.ActionEvent evt) private void enablePanel(JPanel panel) private void disablePanel(JPanel panel) private JTree createTree(Vector nodeVector)</pre> |

4.2.2.6 ParamsFrame.java

| precisgui.ParamsFrame |
|--|
| <i>Attributes</i> |
| <pre>private javax.swing.JPanel buttonPanel; private javax.swing.JTextField conjField; private javax.swing.JLabel conjLabel; private javax.swing.JTextField eolField; private javax.swing.JLabel eolLabel; private javax.swing.JTextField nullField; private javax.swing.JTextField nullFieldInMacros; private javax.swing.JLabel nullLabel; private javax.swing.JLabel nullLabelInMacros; private javax.swing.JPanel paramsPanel; private javax.swing.JButton updateButton; // End of variables declaration private Connection connection = null; private Connection infoConnection = null; DBConnectionParams dbConnParams = null; protected String eol = null; protected String conj_word = null; protected String null_word = null; protected String null_word_macros = null;</pre> |
| <i>Operations</i> |
| <pre>public ParamsFrame(DBConnectionParams dbConnParams) private void initComponents() public void loadParams() private void editTemplateButtonActionPerformed(java.awt.event.ActionEvent evt) private void updateButtonActionPerformed(java.awt.event.ActionEvent evt)</pre> |

4.2.2.7 GraphCreator.java



precisgui.M

Attrib

```
private javax.swing.JButton addHAButto
private javax.swing.JButton addVARBut
private javax.swing.JRadioButton allAttr
private javax.swing.JRadioButton allRac
private javax.swing.JButton arityOfButto
private javax.swing.JComboBox arityOfC
private javax.swing.JRadioButton arityO
private javax.swing.ButtonGroup attrGro
private javax.swing.JPanel attrPanel;
private javax.swing.JCheckBox attrTrans
private javax.swing.JButton autoGenBut
private javax.swing.JPanel buttonPanel;
private javax.swing.JButton clearButton;
private javax.swing.JButton deleteTemp
private javax.swing.JLabel delimLabel;
private javax.swing.JButton eolButton;
private javax.swing.JButton equalsButto
private javax.swing.JTextField fixedSize
private javax.swing.JRadioButton fixedS
private javax.swing.JRadioButton haRac
private javax.swing.JRadioButton joinRa
private javax.swing.JTextField lastLineS
private javax.swing.JLabel lastLineSymb
private javax.swing.JButton lessOrEqua
private javax.swing.JButton lessThanBu
private javax.swing.JTextField lineDelim
private javax.swing.ButtonGroup loopGr
private javax.swing.JPanel loopPropsPa
private javax.swing.JSeparator loopSep
private javax.swing.JPanel loopSizePan
private javax.swing.JTextA50rea macroAre
private javax.swing.JScrollPane macroA
private javax.swing.JPanel macroButte
```

precisgui.MacroEditor

Operations

```
public MacroEditor(DBConnectionParams dbConnParams,
                  RelationEditor previousRelationEditor)

public MacroEditor(DBConnectionParams dbConnParams,
                  JoinRelationsEditor previousJoinRelationsEditor)

public MacroEditor(DBConnectionParams dbConnParams)

private void Creator(DBConnectionParams dbConnParams)

private void initComponents()

private void unWrapTextButtonActionPerformed(java.awt.event.ActionEvent evt)

private void wrapTextButtonActionPerformed(java.awt.event.ActionEvent evt)

private void deleteTemplateButtonActionPerformed(java.awt.event.ActionEvent
evt)

private void deleteTemplateButtonAction()

private void saveTemplateButtonActionPerformed(java.awt.event.ActionEvent evt)

private void clearButtonActionPerformed(java.awt.event.ActionEvent evt)

private void eolButtonActionPerformed(java.awt.event.ActionEvent evt)

private void autoGenButtonActionPerformed(java.awt.event.ActionEvent evt)

private void separateLineBoxActionPerformed(java.awt.event.ActionEvent evt)

private void arityOfButtonActionPerformed(java.awt.event.ActionEvent evt)

private void haRadioButtonActionPerformed(java.awt.event.ActionEvent evt)

private void allAttrsRadioButtonActionPerformed(java.awt.event.ActionEvent evt)

private void lessThanButtonActionPerformed(java.awt.event.ActionEvent evt)\

private void lessOrEqualThanButtonActionPerformed(java.awt.event.ActionEvent
evt)

private void equalsButtonActionPerformed(java.awt.event.ActionEvent evt)

private void plusButtonActionPerformed(java.awt.event.ActionEvent evt)

private void addVARButtonActionPerformed(java.awt.event.ActionEvent evt)

private void addHAButtonActionPerformed(java.awt.event.ActionEvent evt)

private void enablePanel(JPanel panel)

private void disablePanel(JPanel panel)

private boolean selectContentFromDB(boolean isTuplesTranslationBoxSelected,
                                   boolean isHARadioButtonSelected,
                                   boolean isALLRadioButtonSelected,
                                   boolean isPartialRadioButtonSelected,
                                   boolean isSingleRadioButtonSelected)
```

4.2.2.9 *Utils.java*

| <i>precisgui.Utils</i> |
|---|
| <i>Attributes</i> |
| private static Utils utils = null; private static boolean isMacroEditorOpen; |
| <i>Operations</i> |
| private Utils() public static Utils getUtils() public static boolean isMacroEditorOpen() public static void setMacroEditorOpen(boolean setter) |

4.2.2.10 *Translator.java*

| <i>precisgui.Translator</i> |
|--|
| <i>Attributes</i> |
| private javax.swing.JTextField commandField; private javax.swing.JLabel commandLabel; private javax.swing.JTextArea resultArea; private javax.swing.JScrollPane resultAreaScrollPane; private javax.swing.JButton translateButton; |
| <i>Operations</i> |
| public Translator() private void initComponents() private void translateButtonActionPerformed(java.awt.event.ActionEvent evt) |

4.2.2.13 *CheckRenderer.java*

| <i>precisgui.utils.CheckRenderer</i> |
|---|
| <i>Attributes</i> |
| protected JCheckBox check; protected TreeLabel label; |
| <i>Operations</i> |
| public CheckRenderer() public Component getTreeCellRendererComponent(JTree tree, Object value, boolean isSelected, boolean expanded, boolean leaf, int row, boolean hasFocus) public Dimension getPreferredSize() public void doLayout() public void setBackground(Color color) |

4.2.2.14 DBConnectionParams.java

| precisgui.DBConnectionParams |
|---|
| <i>Attributes</i> |
| <pre>private Connection connection; private Connection infoConnection; private String url; private String connDB; private String userName; private String password; private String reposUrl; private String reposDB; private String reposUserName; private String reposPassword; private String profileUserName; private String profilePassword; private String profileID; private String initialRelation; private boolean readOnly = false;</pre> |
| <i>Operations</i> |
| <pre>public DBConnectionParams(String url, String userName, String password, String reposUrl, String reposUserName, String reposPassword, String profileUserName, String profilePassword) throws Exception public void setParams(String url, String userName, String password, String reposUrl, String reposUserName, String reposPassword, String profileUserName, String profilePassword) public void setUrl(String url) public String getUrl() public void setConnDB public String getConnDB() public void setUserName(String userName) public String getUserName() public void setPassword(String password) public String getPassword() public void setReposUrl(String reposUrl) public String getReposUrl() public void setReposDB(String connDB) public String getReposDB() public void setReposUserName(String reposUserName) public String getReposUserName() public void setReposPassword(String reposPassword) public String getReposPassword() public void setProfileUserName(String profileUserName) public String getProfileUserName() public void setProfilePassword(String profilePassword) public String getProfilePassword() public Connection getConnection() public Connection getReposConnection() public String getProfileID() public String getInitialRelation() public boolean getReadOnly()</pre> |

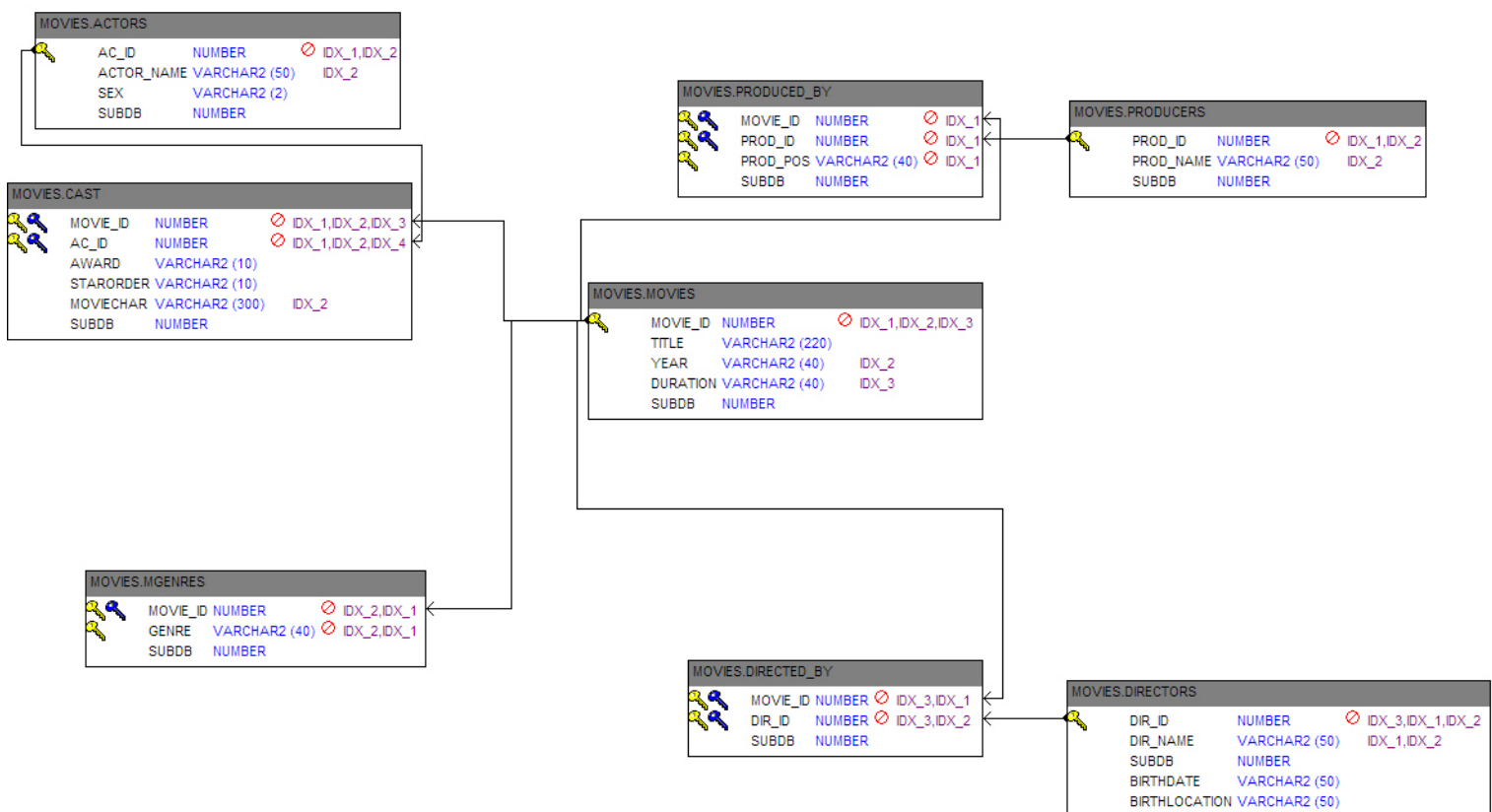
4.3 Βάσεις δεδομένων

Για την λειτουργία της μηχανής μετάφρασης και της γραφικής διεπαφής χρήστη, επικοινωνία με δυο διαφορετικές βάσεις δεδομένων χρειάζεται. Η βάση δεδομένων η οποία προορίζεται για μετάφραση (στην παρούσα περίπτωση η Internet Movie Database), καθώς και μια βάση δεδομένων στην οποία αποθηκεύονται πληροφορίες σχετικές με την μετάφραση, όπως οι ετικέτες των ακμών, οι μακροεντολές και ούτω καθ' εξής.

Το σχήμα της IMDB φαίνεται στο σχήμα 4.1.3.1

Υπάρχει ένα επιπλέον πεδίο, το οποίο ονομάζεται SUBDB, και το οποίο δημιουργήθηκε στην παρούσα περίπτωση για να έχουμε διαφορετικές βάσεις δεδομένων στο ίδιο σχήμα για μεγαλύτερη ευελιξία και ποικιλία στα αποτελέσματα.

Το σχήμα της βάσης – repository είναι στο σχήμα 4.1.3.2 . Δεν υπάρχουν συνδέσεις μεταξύ των σχέσεων γιατί ο σκοπός της βάσης αυτής είναι η αποθήκευση δεδομένων σχετικών με την μετάφραση της άλλης βάσης.



Σχήμα 4.1.3.1 Το σχήμα της IMDB

Σε όλους τους πίνακες έχει προστεθεί μια στήλη SUBDB, η οποία δεν υπήρχε στην αρχική βάση για την ευκολία μας να έχουμε πολλές βάσεις σε μια και να μην χρειάζεται, για να έχουμε ποικιλία δεδομένων και αποτελεσμάτων, να δημιουργούμε ξεχωριστά σχήματα για κάθε υπο- βάση.

| MOVIES_INFO.GRAPH_DIRECTION | | |
|-----------------------------|---------------|---|
| GRAPH_ID | NUMBER | ⊗ |
| INITIAL_RELATION | VARCHAR2 (50) | ⊗ |
| PARENTTAB | VARCHAR2 (50) | ⊗ |
| PARENTCOL | VARCHAR2 (50) | ⊗ |
| CHILDTAB | VARCHAR2 (50) | ⊗ |
| CHILDCOL | VARCHAR2 (50) | ⊗ |

| MOVIES_INFO.USERS | | |
|-------------------|---------------|---|
| USERNAME | VARCHAR2 (50) | ⊗ |
| PASSWORD | VARCHAR2 (50) | ⊗ |
| PROFILE_ID | NUMBER | ⊗ |

| MOVIES_INFO.JOIN_INFO | | |
|-----------------------|----------------|----------------------|
| ID | NUMBER | ⊗ ID _X _1 |
| DB_NAME_P | VARCHAR2 (50) | ⊗ |
| REL_NAME_P | VARCHAR2 (50) | ⊗ |
| DB_NAME_C | VARCHAR2 (50) | ⊗ |
| REL_NAME_C | VARCHAR2 (50) | ⊗ |
| LABEL | VARCHAR2 (150) | ⊗ |
| PROFILE_ID | NUMBER | |
| LIST_NAME_C | VARCHAR2 (30) | |
| WEIGHT | FLOAT | |
| OVERRIDE | VARCHAR2 (1) | |

| MOVIES_INFO.PROJECTION_INFO | | |
|-----------------------------|----------------|----------------------|
| ID | NUMBER | ⊗ ID _X _1 |
| DB_NAME | VARCHAR2 (50) | ⊗ |
| REL_NAME | VARCHAR2 (50) | ⊗ |
| ATTR_NAME | VARCHAR2 (50) | ⊗ |
| LABEL | VARCHAR2 (200) | ⊗ |
| PROFILE_ID | NUMBER | |
| WEIGHT | FLOAT | |

| MOVIES_INFO.LISTS | | |
|-------------------|----------------|---|
| LIST_NAME | VARCHAR2 (50) | ⊗ |
| CONTENT | VARCHAR2 (500) | ⊗ |
| TABLE_NAME | VARCHAR2 (50) | ⊗ |
| LIST_TYPE | VARCHAR2 (20) | ⊗ |
| ENABLED | VARCHAR2 (10) | ⊗ |
| PROFILE_ID | NUMBER | ⊗ |
| LIMIT | VARCHAR2 (20) | ⊗ |

| MOVIES_INFO.PARAMS | | |
|--------------------|---------------|---|
| NAME | VARCHAR2 (50) | ⊗ |
| VALUE | VARCHAR2 (50) | ⊗ |
| PROFILE_ID | NUMBER | ⊗ |

Σχήμα 4.1.3.2 Το σχήμα της βάσης – repository

- GRAPH_DIRECTION :

Χρησιμοποιείται για να υποδείξει ποιο μονοπάτι μετάφρασης θα ακολουθήσει ο αλγόριθμος. Ο χρήστης μπορεί να ορίσει πολλά διαφορετικά μονοπάτια, με διαφορετική αρχική σχέση (initial relation) για το καθένα.

- *GRAPH_ID* : ένας μοναδικός αριθμός- κλειδί που ορίζει το μονοπάτι
- *INITIAL_RELATION* : η αρχική σχέση του μονοπατιού
- *PARENTTAB* : η σχέση- πατέρας
- *PARENTCOL* : το πεδίο της σχέσης- πατέρα που συνδέεται με τη σχέση- παιδί
- *CHILDTAB* : η σχέση- παιδί
- *CHILDCOL* : το πεδίο της σχέσης- παιδί που συνδέεται με τη σχέση- πατέρα

- USERS :

Χρησιμοποιείται για την εισαγωγή διαφορετικών χρηστών στο σύστημα.

Επιτρέπονται διαφορετικά προφίλ χρηστών και κάθε χρήστης μπορεί να δημιουργήσει δικά του μονοπάτια διάσχισης, δικές του ετικέτες και ο,τι άλλο χρειάζεται η μηχανή μετάφρασης. Επίσης, δίνεται η δυνατότητα να επισκοπήσει (αλλά όχι να μεταβάλλει) επιλογές άλλων χρηστών (απλά εισαγόμενος στο σύστημα με το όνομα χρήστη του χρήστη που επιθυμεί να επισκοπήσει).

- *USERNAME* : το όνομα χρήστη για την εισαγωγή στο σύστημα
- *PASSWORD* : ο κωδικός του εκάστοτε χρήστη
- *PROFILE_ID* : ο κωδικός του προφίλ του χρήστη

- *JOIN_INFO* :

Περιέχει όλες τις πληροφορίες για τις ετικέτες και την μετάφραση των ακμών σύνδεσης μεταξύ των κόμβων- σχέσεων.

- *ID* : ένας μοναδικός αριθμός που χαρακτηρίζει την ετικέτα αυτή
- *DB_NAME_P* : η βάση στην οποία βρίσκεται η σχέση- πατέρα
- *REL_NAME_P* : το όνομα της σχέσης- πατέρα
- *DB_NAME_C* : η βάση στην οποία βρίσκεται η σχέση- παιδί (ίδια με την σχέση- πατέρα, η πληροφορία υπάρχει για περαιτέρω ανάπτυξη της εργασίας)
- *REL_NAME_C* : το όνομα της σχέσης- παιδί
- *LABEL* : η ετικέτα της ακμής
- *PROFILE_ID* : ο αριθμός προφίλ του χρήστη
- *LIST_NAME_C* : η μακροεντολή (λίστα) που ορίζει την σχέση- παιδί
- *WEIGHT* : το βάρος της ακμής που ορίζει αν θα μεταφραστεί ή όχι η ακμή στο μονοπάτι της μετάφρασης.
- *OVERRIDE* : υποδεικνύει αν η σχέση είναι για να αποκλειστεί μετάφρασης ή όχι. Ο αριθμός 1 σημαίνει πως δεν θα μεταφραστεί και θα χρησιμοποιηθεί απλά ως συνδετική σχέση, ενώ 0 σημαίνει πως θα μεταφραστεί κανονικά*.

- *PROJECTION_INFO* :

Περιέχει όλες τις πληροφορίες για τις ετικέτες και την μετάφραση των ακμών μεταξύ μιας σχέσης και των κόμβων- πεδίων της.

- *ID* : ένας μοναδικός αριθμός που χαρακτηρίζει την ετικέτα αυτή

* Μόνο για το unary module. Περισσότερες λεπτομέρειες στο κεφάλαιο υλοποίησης.

- *DB_NAME* : η βάση στην οποία βρίσκεται η σχέση και τα πεδία της
 - *REL_NAME* : το όνομα της σχέσης
 - *ATTR_NAME* : το όνομα του πεδίου στο οποίο αναφέρεται η ετικέτα
 - *LABEL* : το περιεχόμενο της ετικέτας (που μπορεί να περιέχει @ΠΕΔΙΑ της σχέσης, εκφράσεις, σχετικά με το @HA - πρωτεύον πεδίο -)
 - *PROFILE_ID* : ο αριθμός προφίλ του χρήστη
 - *WEIGHT* : το βάρος της ακμής που ορίζει αν θα μεταφραστεί ή όχι η ακμή στο μονοπάτι της μετάφρασης.
- *LISTS* :
Περιέχει τις μακροεντολές- λίστες που χρησιμοποιούνται από το σύστημα για την μετάφραση μιας σχέσης ή κάποιων πεδίων μιας σχέσης σε μια ακμή σύνδεσης.
 - *LIST_NAME* : το όνομα της μακροεντολής- λίστας
 - *CONTENT* : το περιεχόμενο της μακροεντολής
 - *TABLE_NAME* : το όνομα της σχέσης στην οποία αναφέρεται η μακροεντολή
 - *LIST_TYPE* : ο τύπος της μακροεντολής (αν περιέχει μετα- δεδομένα της σχέσης, μετάφραση μιας ολόκληρης ή μερικής σχέσης, ή μετάφραση σχέσης μέσα από μια ακμή σύνδεσης είτε ολική είτε μερική είτε μόνο το πρωτεύον πεδίο)
 - *ENABLED* : σε περίπτωση που αναφέρεται στην μετάφραση μιας σχέσης (TR – translation of a relation) και όχι σε ακμή σύνδεσης, εδώ αναφέρεται αν είναι ενεργοποιημένη, ούτως ώστε να αγνοηθούν οι ακμές προβολής που είναι δηλωμένες για την συγκεκριμένη σχέση και υπάρχουν μέσα στην ενεργοποιημένη μακροεντολή
 - *PROFILE_ID* : ο αριθμός προφίλ του χρήστη
 - *LIMIT* : το όριο πάνω από το οποίο δεν θα μεταφραστούν πλειάδες πίνακα. Αυτό έχει επίπτωση και στις επόμενες σχέσεις στο μονοπάτι. Όσες πλειάδες μεταφραστούν από την παρούσα σχέση, μόνο και μόνο αυτές θα συνδεθούν με τις υπόλοιπες σχέσεις στο μονοπάτι της μετάφρασης. Με αυτόν τον τρόπο μπορούμε να μειώσουμε δραστικά το μέγεθος της μετάφρασης, αν το επιθυμούμε (ένας άλλος τρόπος είναι μέσω των βαρών στις ακμές).
 - *PARAMS* :
Περιέχει κάποιες παραμέτρους- συστήματος που μπορεί κάθε χρήστης να μεταβάλλει

όσον αφορά την μετάφραση. Μέχρι τώρα έχουν κριθεί χρήσιμες οι παράμετροι EOL (πώς ορίζεται μια νέα γραμμή στις ετικέτες – στην μετάφραση πώς αναγνωρίζει η μηχανή να αλλάξει γραμμή μέσα από τις ετικέτες. Αρχική τιμή είναι η “\n”), CONJUNCTIVE_WORD (η λέξη- σύνδεσης μεταξύ των ετικετών που χρησιμοποιούνται σε ένα split- module), NULL_VALUES (τι θα βάζει η μηχανή μετάφρασης στο κείμενο μετάφρασης ότι βρίσκει κενή τιμή στην βάση μέσα στις ετικέτες – αν οριστεί κενό τότε η μηχανή μετάφρασης απορρίπτει από την μετάφραση ολόκληρη την πρόταση που περιείχε την κενή τιμή – περισσότερες λεπτομέρειες στα παραδείγματα στο κεφάλαιο 6) και NULL_VALUES_MACRO (το ίδιο με το προηγούμενο, αλλά όταν βρίσκει κενές τιμές εντός μακροεντολών).

- *NAME* : το όνομα της παραμέτρου
- *VALUE* : η τιμή της παραμέτρου
- *PROFILE_ID* : ο αριθμός προφίλ του χρήστη

5

Υλοποίηση

Σε αυτό το κεφάλαιο θα παρουσιασθούν λεπτομερώς θέματα υλοποίησης του συστήματος. Περισσότερο βάρος δίνεται στην μηχανή μετάφρασης, η οποία έχει περισσότερο αλγοριθμικό και τεχνικό ενδιαφέρον, καθώς περιέχει όλη την μεθοδολογία και τους αλγόριθμους μετάφρασης των ερωτήσεων *précis* που αναπτύχθηκαν στην παρούσα εργασία.

5.1 Λεπτομέρειες υλοποίησης

Τα θέματα που ακολουθούν αφορούν κυρίως αλγόριθμους και εφαρμογές της μηχανής μετάφρασης που χρησιμοποιεί για την μετάφραση και διαμόρφωση του κειμένου που επιστρέφεται στον χρήστη. Είναι αφηρημένες εκδόσεις γραμμένες σε ψευδοκώδικα και συνοψίζουν τις λειτουργίες δημιουργίας και εμπλουτισμού του γράφου, εύρεσης κοινών εκφράσεων για ακμές προβολής αλλά και ακμές σύνδεσης σχέσεων, μετάφρασης μιας σχέσης, μετάφρασης σύνδεσης σχέσεων *unary* και *split* τεχνοτροπίας και της λειτουργίας της γραμματικής ανάλυσης και μετάφρασης μακροεντολών.

5.1.1 Υλοποίηση γράφου (κόμβων και ακμών)

Για την μηχανή μετάφρασης, όπως αναφέρεται και παρακάτω, εκτός των άλλων, η βιβλιοθήκη LEDA χρησιμοποιήθηκε, λόγω των δυνατοτήτων που προσφέρει για γράφους. Για την δική μας υλοποίηση γράφου όμως με τα συγκεκριμένα χαρακτηριστικά των κόμβων των ακμών, δημιουργήσαμε δυο κλάσεις που περιέχουν τα χαρακτηριστικά που επιθυμούμε.

Για την κλάση των κόμβων :

5.1.1.1 Κόμβοι

```
01. class aNode {
02.     public:
03.         // attributes
04.         leda_node n;
05.         int id;
06.         leda_string name;
07.         leda_string type;
08.         int heading;
09.         int initial_relation;
10.         leda_list<leda_string> translated;
11.         leda_string pkey;
12.         int in;
13.         int out;

14.         // methods
15.         void Print();
16.         aNode() :
17.             id(-1), n(nil), name(), type(), heading(),
18.             weight 0.0), label(), visited(), initial_relation() {}
19.         aNode(const aNode& ND) {
20.             id=ND.id; n=ND.n; name=ND.name; type=ND.type;
21.             heading=ND.heading; visited=ND.visited;
22.             weight=ND.weight; label=ND.label;
23.             initial_relation=ND.initial_relation; in=ND.in;
24.             out=ND.out; translated=ND.translated; pkey=ND.pkey;
25.         }
26.         aNode& operator = ( const aNode& ND ){
27.             id=ND.id; n=ND.n; name=ND.name; type=ND.type;
28.             heading=ND.heading; weight=ND.weight; label=ND.label;
29.             visited=ND.visited; initial_relation=ND.initial_relation;
30.             in=ND.in; out=ND.out; translated=ND.translated;
31.             pkey=ND.pkey;
32.             return *this;
```

```
33.     }
34.   };
```

Για τα attributes :

leda_node n;

Είναι ο κόμβος γράφου που αναγνωρίζει η βιβλιοθήκη LEDA.

int id;

Για κάθε κόμβο, υπάρχει ένας μοναδιαίος αριθμός που υποδηλώνει την ταυτότητα του κόμβου.

leda_string name;

Εδώ αποθηκεύεται το όνομα του κόμβου (όνομα σχέσης αν πρόκειται για σχέση, ή πεδίου αντίστοιχα αν πρόκειται για πεδίο κάποιας σχέσης).

leda_string type;

Αυτό το πεδίο διαφοροποιεί τους κόμβους σχέσης ή πεδίου. Αν πρόκειται για κόμβο σχέσης τότε έχει την τιμή “relation” και αν πρόκειται για πεδίο, έχει την τιμή “attribute”.

int heading;

Αν πρόκειται για κόμβο “attribute”, τότε αυτός ο αριθμός είναι 1 αν το πεδίο αυτό είναι πρωτεύον, ή 0 σε διαφορετική περίπτωση.

int initial_relation;

Αν πρόκειται για κόμβο “relation”, αν η σχέση είναι αρχική, τότε αυτός ο αριθμός είναι 1, ή 0 σε κάθε άλλη περίπτωση.

leda_list<leda_string> translated;

Αναφέρεται σε κόμβους σχέσης και είναι μια λίστα από πρωτεύοντα κλειδιά της σχέσης αυτής που έχουν ήδη μεταφραστεί, έτσι ώστε αν εμφανιστεί πάλι για μετάφραση η συγκεκριμένη πλειάδα που αναφέρεται σε άποιο πρωτεύον κλειδί που περιέχεται στην λίστα, η μετάφρασή του θα περιλαμβάνει μόνο μετάφραση του πρωτεύοντος πεδίου (σύμφωνα με την σημειογραφία μας, μετάφραση R_0 αντί για R_x).

leda_string pkey;

Πάλι, αναφέρεται σε κόμβους σχέσης και είναι το όνομα του πεδίου του πρωτεύοντος κλειδιού για την συγκεκριμένη σχέση.

int in;

Για κόμβους σχέσης, ο αριθμός των εισερχόμενων ακμών.

int out;

Για κόμβους σχέσης, ο αριθμός των εξερχόμενων ακμών.

Τα δυο τελευταία πεδία χρησιμοποιούνται από τον αλγόριθμο TRLS (αναλύεται παρακάτω) για να αποφασίσει αν θα χρησιμοποιήσει unary ή split module.

Για τις μεθόδους (methods) :

void Print();

Μια μέθοδος για την όμορφη εμφάνιση δεδομένων των κομβων του γράφου στην οθόνη. Ο κώδικάς της δεν εμφανίζεται εδώ, βρίσκεται στο header file της κλάσης.

aNode()..

aNode(const aNode& ND)..

Δυο κατασκευαστές (constructors) της κλάσης.

aNode& operator = (const aNode& ND)..

Υπερφόρτωση (overload) του τελεστή (operator) “=”.

5.1.1.2 Ακμές

```
01. class anEdge {
02.     public:
03.     //attributes
```

```

04.   leda_edge e;
05.   int id;
06.   leda_string name;
07.   leda_string type;
08.   leda_string label;
09.   leda_string join_p_key;
10.   leda_string join_c_key;
11.   double weight;
12.   // methods
13.   void Print();
14.   anEdge():
15.       id(-1), e(nil), name(), type(),
16.       weight(0.0), join_p_key(), join_c_key(),
17.       label(), primary() {}
18.   anEdge(const anEdge& ED) {
19.       id=ED.id; e=ED.e; name=ED.name; type=ED.type;
20.       weight=ED.weight; label=ED.label; primary=ED.primary;
21.       join_p_key=ED.join_p_key; join_c_key=ED.join_c_key;
22.   }
23.   anEdge& operator = ( const anEdge& ED ){
24.       id=ED.id; e=ED.e; name=ED.name; type=ED.type;
25.       weight=ED.weight; label=ED.label; primary=ED.primary;
26.       join_p_key=ED.join_p_key; join_c_key=ED.join_c_key;
27.       return *this;
28.   }
19.   };

```

Για τα attributes :

leda_edge e;

Είναι η ακμή γράφου που αναγνωρίζει η βιβλιοθήκη LEDA.

int id;

Για κάθε ακμή, υπάρχει ένας μοναδιαίος αριθμός που υποδηλώνει την ταυτότητα της ακμής.

leda_string name;

Εδώ αποθηκεύεται το όνομα της ακμής το οποίο ορίζουμε να είναι το όνομα του πρωτεύοντος κλειδιού που συνδέει τις δυο σχέσεις, αν πρόκειται για ακμή σύνδεσης σχέσεων.

leda_string type;

Αυτό το πεδίο διαφοροποιεί τις ακμές σύνδεσης σχέσεων ή ακμές προβολής. Αν πρόκειται για ακμή που συνδέει δυο σχέσεις τότε έχει την τιμή “join” και αν πρόκειται ακμή προβολής έχει την τιμή “projection”.

leda_string label;

Εδώ κρατάμε την ετικέτα που περιέχει απαραίτητα στοιχεία μετάφρασης. Εδώ φορτώνονται δεδομένα από την βάση-repository τα οποία φτιάχνει ο χρήστης με το γραφικό εργαλείο (GUI).

leda_string join_p_key;

Πάνω στις ακμές σύνδεσης σχέσεων περιέχονται τα ονόματα των στηλών των πρωτευόντων κλειδιών που είναι αναγκαία για την εύρεση των πλειάδων που αντιστοιχούν από την σχέση-πατέρα προς την σχέση-υιο. Εδώ περιέχεται το όνομα του πρωτεύοντος κλειδιού της σχέσης-πατέρα.

leda_string join_c_key;

Κατ’ αντιστοιχία με το παραπάνω, εδώ περιέχεται το όνομα του πρωτεύοντος κλειδιού της σχέσης-παιδιού.

Για τις μεθόδους (methods) :

void Print();

Μια μέθοδος για την όμορφη εμφάνιση δεδομένων των ακμών του γράφου στην οθόνη. Ο κώδικάς της δεν εμφανίζεται εδώ, βρίσκεται στο header file της κλάσης.

anEdge()..

anEdge(const anEdge& ED)..

Δυο κατασκευαστές (constructors) της κλάσης..

anEdge& operator = (const anEdge& ED)..

Υπερφόρτωση (overload) του τελεστή (operator) “=”.

5.1.2 *Algorithm Resolve Common Expressions*

Ενδιαφέρον επίσης παρουσιάζει ο μηχανισμός εύρεσης κοινών εκφράσεων (resolve common expressions) σε ετικέτες, είτε αυτές είναι από ακμές προβολής (σε κάθε μετάφραση σχέσης ενεργοποιείται αυτός ο μηχανισμός), είτε είναι ακμές σύνδεσης σχέσεων (ο μηχανισμός για αυτήν την λειτουργία ενεργοποιείται στο split module της μετάφρασης).

Η λογική του αλγορίθμου είναι να βρει κοινές λέξεις/εκφράσεις ανάμεσα σε προτάσεις που πρόκειται να ομαδοποιηθούν. Αυτό συμβαίνει σε κάθε μετάφραση σχέσης (R^\bullet και όχι R°), για την οποία χρησιμοποιούνται τόσες προτάσεις, όσα πεδία έχει η σχέση εκτός από ένα (το πρωτεύον πεδίο). Έτσι, αντί να έχουμε ένα πλήθος προτάσεων $N-1$, όπου N είναι ο αριθμός των πεδίων της σχέσης, καταλήγουμε, ανάλογα με το αν έχουν κοινές εκφράσεις οι προτάσεις, με έναν αριθμό προτάσεων $1..N-1$. Το ίδιο συμβαίνει και για το split module, όπου, αντί να συγκρίνουμε προτάσεις στις ετικέτες των ακμών προβολής, συγκρίνουμε προτάσεις στις ετικέτες των ακμών σύνδεσης σχέσεων.

Ακολουθεί ψευδοκώδικας που επεξηγεί την λειτουργία των δυο διαφορετικών αυτών λειτουργιών. Βρίσκονται στην ίδια συνάρτηση, με αμελητέες διαφοροποιήσεις. Για τον λόγο αυτό, ο ακόλουθος κώδικας παρουσιάζει την ιδέα και την γενική υλοποίηση.

```
01. Input : a set of labels
02. Output : a set of labels, without common expressions
03. Begin
04. result = first_label;
05.     for (every other label) {
```

```

06.     labelTokensList = split the tokens in the label;
07.     result_token = first_token_of_result;
08.     for (every token in labelTokensList){
09.         if(token and result_token are expressions){
10.             exprTokensList = split the token;
11.             resultExprTokensList = split the result_token;
12.             for (every word tokens in both lists){
13.                 if(words are equal){
14.                     delete token from exprTokensList;
15.                 }
16.                 else{
17.                     exit inner and outer loop;
18.                 }
19.             }
20.         }
21.         else if (token == variable){
22.             delete token from labelTokensList;
23.         }
24.         else {
25.             break;
26.         }
27.         result_token = next_token_of_result;
28.     }
29.     connect remaining labelTokensList to result;
30. }
31. return result;
32. End

```

Βάλε στο αποτέλεσμα μια ετικέτα (04.). Για κάθε άλλη ετικέτα (05.), κάνε τα εξής: χώρισε την ετικέτα στα σύμβολα που την αποτελούν (05. - 07. είτε είναι εκφράσεις “..”, είτε είναι μεταβλητές @ΠΕΔΙΟ, ΜΑΚΡΟΕΝΤΟΛΗ). Πάρε το πρώτο σύμβολο του αποτελέσματος και σύγκρινέ το με το πρώτο σύμβολο της ετικέτας. Αν και τα δυο είναι εκφράσεις αλφαριθμητικών, χώρισέ τις στις λέξεις που τις αποτελούν (09. - 11.). Για κάθε λέξη στις δυο

λίστες, όσο είναι ίδιες, σβήνε την λέξη στην ετικέτα που εξετάζουμε (12. - 15.). Αν βρούμε διαφορετικές λέξεις, τότε σύνδεσε στο αποτέλεσμα την υπόλοιπη ετικέτα που έχει απομείνει με σημεία στίξης αν πρόκειται για μετάφραση σε ακμές προβολής, ή με τις συνδετικές λέξεις που έχει ορίσει ο χρήστης για τις ακμές σύνδεσης σχέσεων και προχώρα με την επόμενη ετικέτα (16. - 18., 29., 05.). Αν τελειώσει το αλφαριθμητικό και τα δυο είναι ίδια, πάρε το επόμενο σύμβολο της ετικέτας (27.). Σε περίπτωση που τα σύμβολα στο αποτέλεσμα και την ετικέτα είναι μεταβλητές, αν είναι ίδια προχώρα με το επόμενο σύμβολο (21. - 23., 27.). Αλλιώς, πάρε την επόμενη ετικέτα (05.) και κάνε τα ίδια βήματα.

Για παράδειγμα, ας θεωρήσουμε την σχέση DIRECTORS στην βάση μας. Ας θεωρήσουμε ότι στην μετάφραση της σχέσης, οι ακμές προβολής BIRTHDATE, BIRTHLOCATION χρησιμοποιούνται. Ας θεωρήσουμε ότι έχουν τις εξής τιμές :

Ετικέτα (h_r, BIRTHDATE) = @DIR_NAME+"was born on"+@BIRTHDATE

Ετικέτα (h_r, BIRTHLOCATION) = @DIR_NAME+" was born in"+@BIRTHLOCATION

Τα βήματα που θα ακολουθήσει ο αλγόριθμος θα είναι :

```
01) result = @DIR_NAME+"was born on"+@BIRTHDATE
    processedLabel = @DIR_NAME+" was born in"+@BIRTHLOCATION
```

```
02) result = @DIR_NAME+"was born on"+@BIRTHDATE
    processedLabel = " was born in"+@BIRTHLOCATION
```

```
03) result = @DIR_NAME+"was born on"+@BIRTHDATE
    processedLabel = " born in"+@BIRTHLOCATION
```

```
04) result = @DIR_NAME+"was born on"+@BIRTHDATE
    processedLabel = " in"+@BIRTHLOCATION
```

```
05) result = @DIR_NAME+"was born on"+@BIRTHDATE
    processedLabel = " in"+@BIRTHLOCATION
```

```
06) result = @DIR_NAME+"was born on"+@BIRTHDATE
    +", "+"in"+@BIRTHLOCATION
```

Ίδια επεξεργασία πραγματοποιείται και για τις ετικέτες σε ακμές σύνδεσης σχέσεων, με την διαφορά ότι συνδυαστικοί κρίκοι των προτάσεων είναι λέξεις ορισμένες από τον χρήστη και όχι απλά σημεία στίξης, εδώ.

5.1.3 Algorithm Translation of a Logical Subset (TRLS)

Ακολουθεί ο αλγόριθμος που διατρέχει τον γράφο και αποφασίζει για την μονάδα που θα αναλάβει την μετάφραση κάθε στιγμή (unary ή split).

Μετάφραση λογικού υποσυνόλου σχέσεων :

```

01. Input : a logical subset L, constraint T
02. Output : an array of phrases précis
03. Begin
04. Foreach expanded subgraph  $S \in L$  {
05.     précis [S] = ''
06.     Start from the root  $R_{SG}$  of the respective initial subgraph
                                                     $S_G$ 
07.     Foreach term  $t \in R_{SG}$  {
08.          $QP \leftarrow \{R_{SG}\}$ 
09.         While (QP not empty) {
10.             pick head  $R_i$  of QP
11.             If  $\text{in-deg}(R_i) \leq 1$  {
12.                  $QP \leftarrow \text{find the neighboring relations } R_{j1}, \dots, R_{jk}$ 
13.                 If  $\text{out-deg}(R_i) = 1$ 
14.                     {evaluate  $M_U(R_i, R_j)$ }
15.                 If  $\text{out-deg}(R_i) > 1$ 
16.                     {evaluate  $M_S(R_i, R_{j1}, \dots, R_{jk})$ }
17.             }
18.             else
19.                 {evaluate  $M_U(R_{j1}, R_{j2} \dots, R_i)$ }
20.             continue with next  $R_i$ 
21.             update précis [S]
22.         }

```

23. }
24. Return

Για κάθε ακολουθία κόμβων που είναι δηλωμένη στην βάση – repository (04. μια κάθε φορά), ξεκίνα την μετάφραση με κενό κείμενο (05.). Πρώτος κόμβος που επισκέπτεται είναι η αρχική σχέση που είναι δηλωμένη στην βάση (06.). Για κάθε όρο των περιορισμών του χρήστη (07., εδώ υπάρχει μια λέξη-κλειδί και υπάρχει στον υπογράφο που εξετάζουμε), προσέθεσε στην λίστα την αρχική σχέση (08.). Όσο η λίστα δεν είναι κενή, πάρε το πρώτο στοιχείο-κόμβο της λίστας και βρες τον αριθμό των των εισερχομένων ακμών σύνδεσης (10.). Αν ο αυτός είναι το πολύ 1 (11.), τότε βρες τις γειτονικές σχέσεις (12.). Αν ο αριθμός των εξερχόμενων ακμών σύνδεσης είναι 1, τότε ενεργοποίησε το unary module, αλλιώς αν ο αριθμός των εξερχόμενων ακμών σύνδεσης είναι μεγαλύτερος από 1, τότε ενεργοποίησε το split module (13. – 16.). Αν ο αριθμός των εισερχομένων ακμών σύνδεσης είναι μεγαλύτερος από 1, τότε ενεργοποίησε πάλι το unary module (18.,19.). Ανανέωσε την απάντηση και συνέχισε με το επόμενο στοιχείο της λίστας (20. - 21.).

5.1.4 Algorithm Translation of a Relation (TR)

Ακολουθεί ο αλγόριθμος που επιδεικνύει την μετάφραση μιας σχέσης.

01. Input : a LEDA graph G , a source node N_s ,
02. Output : a string containing translation of the node
and the join with the neighboring node(s)
03. Begin
04. result='';
05. if(macro is defined for this relation){
06. if(macro == metadata list){
07. result = translate metadata;
08. return result;
09. }
10. else if(macro == list){
11. foundList = true;
12. }
13. }
14. $P_E = \{ \text{projection edges that are not in macro} \};$
15. forall(items in P_E){


```

16.  resolvedList = resolveCommonExpressions(PE);
17.  }
18.
19.  AT = { find the Attributes to be translated };
20.
21.  if(foundList = true){
22.      macro = retrieve macro from repository;
23.      result += parse macro;
24.  }
25.  for (each tuple in the source relation){
26.      if(tuple has been translated R* somewhere else){
27.          AT = { Ro translation };
28.          result += Query the DB and replace
29.              the corresponding attributes to AT
30.              replace null values with selection
31.              from the user;
32.      }
33.      else{
34.          result += Query the DB and replace
35.              the corresponding attributes to AT
36.              replace null values with selection
37.              from the user;
38.          append the primary key to the relation,
39.              flagging R* the tuple;
40.      }
41.  }
42.  result = punctuate();
43.  return result;
44.  End

```

Ο αλγόριθμος που μεταφράζει μια σχέση ξεκινάει ελέγχοντας αν υπάρχει κάποια μακροεντολή ορισμένη για αυτήν την σχέση στην βάση-repository. Υπάρχουν δυο ειδών μακροεντολές που αφορούν την μετάφραση μιας σχέσης, οι μακροεντολές που περιέχουν μετα-δεδομένα της βάσης (06. όπως για παράδειγμα τα ονόματα των πεδίων των πινάκων), είτε μακροεντολές που περιέχουν τα πεδία των πινάκων με εκφράσεις (10.). Η δεύτερη

περίπτωση είναι για ευκολία του χρήστη, να μην γράφει ξεχωριστά κάθε ετικέτα για κάθε πεδίο του πίνακα, αν αυτός το επιθυμεί. Στην περίπτωση που η μακροεντολή αφορά μετα-δεδομένα της βάσης, τότε μια μόνο μετάφραση αρκεί, δεν χρειάζεται για κάθε πλειάδα του πίνακα που χρειάζεται μετάφραση να δείχνουμε τα μετα-δεδομένα, αφού είναι ίδια (07. – 08.). Αν, υπάρχει άλλη μακροεντολή, τότε ελέγχουμε ποια πεδία δεν υπάρχουν στην μακροεντολή αυτή και αντίστοιχα, τα προετοιμάζουμε προς μετάφραση (10. – 12.). Δίνεται η δυνατότητα στον χρήστη να αναμίξει τους τύπους μετάφρασης (μακροεντολή και απλή μετάφραση ακμών προβολής), εάν το επιθυμεί. Για τα πεδία που δεν περιέχονται στην μακροεντολή, αν υπάρχει καμία, τρέχει ο αλγόριθμος εύρεσης κοινών εκφράσεων για να κατασκευάσει προτάσεις χωρίς διπλές εκφράσεις (14. – 17.). Έπειτα, επόμενο βήμα είναι να βρούμε τις μεταβλητές-πεδία (@ΠΕΔΙΟ) που βρίσκονται εντός των προτάσεων που είναι προετοιμασμένες για μετάφραση. Βάσει αυτών, θα δημιουργηθεί μια ερώτηση στην βάση για να ανακτήσουμε τα δεδομένα (19.). Αν από τον προηγούμενο έλεγχο μακροεντολής, υπάρχει ενεργοποιημένη μακροεντολή για την σχέση, την ανακτούμε από την βάση-repository και την επεξεργαζόμαστε με τον αλγόριθμο ανάλυσης μακροεντολών για να πάρουμε το αποτέλεσμα-κείμενο (21. – 24.). Έπειτα, για κάθε πλειάδα της σχέσης που χρειάζεται μετάφραση, κατασκευάζουμε μια ερώτηση στην βάση που περιέχει τα πεδία που βρίσκονται στις προτάσεις που επεξεργάστηκε ο αλγόριθμος εύρεσης κοινών εκφράσεων. Αν η πλειάδα της σχέσης έχει μεταφραστεί ως R^* ήδη κάποια προηγούμενη στιγμή (κλασικό παράδειγμα είναι μια σχέση που είναι σχέση-παιδί δυο ή παραπάνω άλλων σχέσεων), τότε μεταφράζεται μόνο σε μορφή R^o (26. – 32.). Αλλιώς, μεταφράζεται σε R^* και μαρκάρεται η πλειάδα αυτή πως έχει μεταφραστεί ως R^* (33. – 40.). Αφού πάρουμε τα αποτελέσματα από την βάση δεδομένων, τα αντικαθιστούμε στις προτάσεις, αντικαθιστώντας όπου έχει <null> τιμές με τιμή, ο,τι έχει δώσει ο χρήστης να φαίνεται στην μετάφραση σε αντίστοιχη περίπτωση. Δίνεται η δυνατότητα, αν ο χρήστης έχει δώσει κενό πεδίο όπου ορίζει την συμπεριφορά της μηχανής μετάφρασης για τις κενές τιμές, να απορριφθούν οι ετικέτες των ακμών αυτές από την μετάφραση τελείως*. Τέλος, το αποτέλεσμα εμπλουτίζεται με σωστά σημεία στίξης και επιστρέφεται (42. – 44.).

*Παράδειγμα :

Ετικέτα (hr, BIRTHDATE) = @DIR_NAME+"was born on"+@BIRTHDATE

Ετικέτα (hr, BIRTHLOCATION) = @DIR_NAME+" was born in"+@BIRTHLOCATION

Παρατηρήσαμε πριν πως το αποτέλεσμα για την μετάφραση της σχέσης DIRECTORS με τις παραπάνω ετικέτες, ήταν :

```
result = @DIR_NAME+"was born on"+@BIRTHDATE
        +", "+"in"+@BIRTHLOCATION
```

Και μετά από την ερώτηση στην βάση, το αποτέλεσμα της μετάφρασης θα ήταν:

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935.

Εάν το πεδίο @ BIRTHDATE ήταν null στην βάση δεδομένων, τότε αν ο χρήστης έχει δηλώσει στην βάση-repository κενό το πεδίο για την αντικατάσταση των κενών τιμών από την βάση, τότε το αποτέλεσμα θα ήταν :

Allen, Woody was born in Brooklyn, New York, USA.

ενώ αν είχε δηλώσει, για παράδειγμα <n/a>, το αποτέλεσμα θα ήταν :

Allen, Woody was born in Brooklyn, New York, USA, on <n/a>.

5.1.5 Algorithm Unary Translation (UTR)

Πιο συγκεκριμένα, ακολουθεί ψευδο-κώδικας που αναλύει την λειτουργία της μετάφρασης unary.

```
01. Input : a LEDA graph G, an edge E, a list containing
02.         the join query strings, and a list of the nodes
03.         to be traversed.
04. Output : a string containing translation of the join
05.         of the two nodes connected by the edge
06. Begin
07. result='';
08. Foreach(tuple of source relation){
09.     label = label of the edge;
10.     macro = retrieve and parse the macro
11.     corresponding to the target relation;
12.     result += Query the DB and replace
13.             the corresponding attributes and
14.             macros to the label
15.             replace null values with
16.             selection from the user;
17.     joinString = generate next relation joinString;
```

```

18.         global_wheres.append(joinString);
19.         global_nodes.append(target_node);
20.     }
21.     result = punctuate();
22.     return result;
23. End

```

Αμέσως μετά την μετάφραση κάποιας σχέσης, λαμβάνει χώρα η μετάφραση της σύνδεσης αυτής της σχέσης με τις γειτονικές της (αν είναι τελική σχέση, η μετάφραση σταματάει εκεί). Στην περίπτωση του unary join, όπως αυτό έχει οριστεί και στο θεωρητικό κομμάτι, για κάθε πλειάδα της σχέσης, αρχικοποιούμε την ετικέτα της σύνδεσης των σχέσεων (08. – 09.). Κάθε ετικέτα που βρίσκεται σε μια ακμή σύνδεσης περιέχει το πρωτεύον πεδίο της σχέσης-πατέρα, κάποιες εκφράσεις αλφαριθμητικών και μια μακροεντολή που αναφέρεται στην σχέση-παιδί. Ο αλγόριθμος ανακτά από την βάση- repository την μακροεντολή αυτή και τρέχει τον αλγόριθμο ανάλυσης μακροεντολών που επιστρέφει το αποτέλεσμα- κείμενο (10.). Επόμενο βήμα είναι η δημιουργία ερώτησης προς την βάση που περιλαμβάνει το πρωτεύον πεδίο της σχέσης-πατέρα και η αντικατάσταση της μεταβλητής με το περιεχόμενο από την βάση (12. – 16.).

Στην μηχανή μετάφρασης, εκτός των άλλων, υπάρχουν δυο πολύ σημαντικές λίστες που έχουν παγκόσμια εμβέλεια σε όλον τον κώδικα και που περιέχουν

α)κάποια αλφαριθμητικά που παρέχουν τους περιορισμούς για την σύνδεση των σχέσεων (18. ποια πρωτεύοντα κλειδιά θα συνδέσει από την σχέση-πατέρα και σε ποια κλειδιά αντιστοιχούν στην σχέση-παιδί)

β) τους κόμβους που πρόκειται να μεταφραστούν στα επόμενα βήματα (19.).

Τέλος, δημιουργούνται οι περιορισμοί για την σύνδεση με την επόμενη σχέση και εισέρχονται στην λίστα, καθώς και η σχέση-παιδί στην λίστα με τους επόμενους κόμβους και επιστρέφεται το αποτέλεσμα (17. – 19.).

5.1.6 Algorithm Split Translation (STR)

Πιο συγκεκριμένα, ακολουθεί ψευδο-κώδικας που αναλύει την λειτουργία της μετάφρασης split.

```

01. Input   : a LEDA graph G, a set of edges Es{..},
02.         a list containing the join query strings, and

```

```

03.         a list of the nodes to be traversed.
04. Output : a string containing translation of the join
05.         of the nodes involved
06. Begin
07. result='';
08. forall(edge in Es){
09.         resolvedList = resolveCommonExpressions(Es);
10.         label = resolvedList.toString();
11. }
12. Foreach(tuple of source relation){
13.     Foreach(edge in Es){
14.         macro = retrieve and parse the macro;
15.             corresponding to each target relation;
16.         joinString = generate next relation joinString;
17.         global_wheres.append(joinString);
18.         global_nodes.append(target_node);
19.     }
20.     result += Query the DB and replace
21.         the corresponding attributes and macros
22.         to the label replace null values with
23.         selection from the user;
24. }
25. result = punctuate();
26. return result;
27. End

```

Στον αλγόριθμο αυτό, ο οποίος μοιάζει σε αυτήν την απλοποιημένη έκδοση με ψευδοκώδικα με την μετάφραση με το unary module, αρχικά μαζεύονται οι ακμές σύνδεσης των σχέσεων-παιδιών και μέσω του αλγορίθμου εύρεσης κοινών εκφράσεων απορρίπτουμε κοινές εκφράσεις των προτάσεων, χάριν ευμορφίας (08. – 11.). Ως αποτέλεσμα, παίρνουμε μια μεγάλη έκφραση που περιέχει την σύνθεση των εμπλεκόμενων ακμών. Επόμενο βήμα είναι να μεταφράσουμε για κάθε πλειάδα της σχέσης- πατέρα, το αποτέλεσμα της μετάφρασης για την παραπάνω σύνθεση. Για κάθε εμπλεκόμενη ακμή, μαζεύουμε πληροφορίες για την μακροεντολή που την προσδιορίζουν και την αντικαθιστούμε στην ετικέτα που έχουμε για να παράγουμε το επιθυμητό αποτέλεσμα (12. – 15.). Επιπλέον, για κάθε ακμή καθορίζουμε

τους περιορισμούς που την προσδιορίζουν (και που προορίζονται για την σύνδεσή της με επόμενες σχέσεις- κόμβους) και τις προσθέτουμε στην αντίστοιχη λίστα με τους περιορισμούς, καθώς και τον κόμβο τον ίδιο στην αντίστοιχη λίστα με τους κόμβους προς επίσκεψη (16. – 18.). Τέλος, αντικαθίστανται οι μεταβλητές πεδίων (@ΠΕΔΙΟ – συνήθως το πρωτεύον πεδίο), καθώς και τα αποτελέσματα από τις μακροεντολές παραπάνω, και έπειτα από την συνήθη επεξεργασία με σημεία στίξης, επιστρέφεται το αποτέλεσμα (20. – 27.).

5.1.7 *Algorithm Translation of a macro*

Ο επόμενος αλγόριθμος παρουσιάζει την λειτουργία της γραμματικής ανάλυσης και μετάφρασης μακροεντολών :

```

01. Input   : a macro in the form of a string, a set of constraints
02.         for the query to the database,
03.         and the type of translation T
04. Output  : a string containing translation of the macro
05. Begin
06. header = get the header of the macro;
07. list_type = parse the header and get
08.           the list type ( 5 types available);
09. Foreach( loop in the header, defined by [...]{} ){
10.     loop_iteration = parse the loop header;
11.     attributes AM = {
12.       parse the loop content and
13.       find the attributes;
14.     }
15.     result += query database, substitute the @VARIABLES,
16.              take care of the null values,
17.              take care of the primary keys Ro, R•,
18.              all according to list_type;
19.   }
20. result = punctuate();
21. return result;
22. End

```

Παράδειγμα μακροεντολής :

DEFINE MOVIES_ATTR_LIST as

[i < arityOf(@MOVIES)]{@MOVIES[\$i\$]+","}

[i = arityOf(@MOVIES)]{@MOVIES[\$i\$]+"."}

Ο παραπάνω αλγόριθμος αποτελεί την γραμματική ανάλυση και μετάφραση των μακροεντολών. Οι μακροεντολές είναι πέντε ειδών από τις οποίες δυο αφορούν την μετάφραση μιας σχέσης – είτε μετα-δεδομένα της βάσης είτε εκφράσεις με μεταβλητές –, και οι άλλες τρεις αφορούν την μετάφραση κάποιας σχέσης, αλλά μέσα από μια ακμή σύνδεσης με μια άλλη σχέση. Αυτές αφορούν την μετάφραση μόνο του πρωτεύοντος πεδίου (R°), μετάφραση όλων ή ενός αριθμού από πεδία (και οι δυο θεωρούνται ως μετάφραση R^\bullet). Ο αλγόριθμος λοιπόν, ξεκινάει πρώτα με την εύρεση της επικεφαλίδας της μακροεντολής που περιέχει πληροφορίες για τον τύπο της μακροεντολής (06. – 08.)

(εδώ : *DEFINE MOVIES_ATTR_LIST as*).

Έπειτα, για κάθε βρόχο της μακροεντολής που ορίζεται από τις αγκύλες [κεφαλίδα βρόχου] { περιεχόμενο βρόχου } (09.)

(εδώ : *[i < arityOf(@MOVIES)]{@MOVIES[\$i\$]+","}*),

αναγνωρίζεται από την κεφαλίδα του βρόχου ο τελεστής και τα όρια του βρόχου – που μπορεί να είναι μια συνάρτηση όπως η *arityOf(@ΠΕΔΙΟ)* ή κάποιος σταθερός αριθμός –, και από το περιεχόμενο του βρόχου λαμβάνονται οι εκφράσεις και οι μεταβλητές από τις οποίες θα λάβουμε το αποτέλεσμα της μετάφρασης (10. – 14.). Ανάλογα τον τύπο της μακροεντολής, κατασκευάζεται η ερώτηση στην βάση δεδομένων και αντικαθίστανται οι μεταβλητές με δεδομένα από την βάση (15. – 18.). Κάθε απάντηση από την βάση επεξεργάζεται ελαφρώς χάριν ευμορφίας και κενά πριν ή μετά του περιεχομένου απορρίπτονται. Αν η μετάφραση δεν περιλαμβάνει μόνο το πρωτεύον πεδίο της σχέσης, τότε εισέρχεται στην αντίστοιχη λίστα του κόμβου- σχέσης το πρωτεύον κλειδί των πλειάδων που μεταφράζονται ως R^\bullet , για να αποφύγουμε επανάληψη της μετάφρασης της συγκεκριμένης πλειάδας πάνω από μια φορές. Τέλος, επεξεργάζεται η απάντηση και προστίθενται όπου χρειάζονται σημεία στίξης και επιστρέφεται το αποτέλεσμα (20. – 22.).

Αναλυτικά παραδείγματα θα ακολουθήσουν στην ενότητα επίδειξης των εργαλείων.

5.2 Πλατφόρμες και προγραμματιστικά εργαλεία

Στο παρόν κεφάλαιο θα γίνει μια αναφορά και σύντομη ανάλυση των προγραμματιστικών εργαλείων και των περιβαλλόντων που χρησιμοποιήθηκαν κατά την διαδικασία της κωδικοποίησης του λογισμικού συστήματος.

5.2.1 Λειτουργικό σύστημα

Ως λειτουργικό σύστημα για την ανάπτυξη και τον έλεγχο των εφαρμογών που δημιουργήθηκαν χάριν αυτής της εργασίας (η μηχανή μετάφρασης και η γραφική διεπαφή του χρήστη με τις πληροφορίες της μετάφρασης που είναι στην βάση δεδομένων) χρησιμοποιήθηκε το Microsoft Windows XP Professional Service Pack 2.

Η χρήση των Windows κρίθηκε αναγκαία για την ανάπτυξη της μηχανής μετάφρασης η οποία αναπτύχθηκε σε Microsoft Visual Studio. Για την γραφική διεπαφή (GUI) με την βάση δεδομένων, η οποία γράφτηκε σε γλώσσα προγραμματισμού Java μπορεί να λειτουργήσει ανεξάρτητα από το λειτουργικό σύστημα.

5.2.2 Γλώσσες προγραμματισμού

Για τις ανάγκες της διπλωματικής, δυο διαφορετικές γλώσσες προγραμματισμού χρησιμοποιήθηκαν για την δημιουργία των δυο εργαλείων που αναπτύχθηκαν, την μηχανή μετάφρασης η οποία αναπτύχθηκε σε γλώσσα προγραμματισμού C++, και η γραφική διεπαφή χρήστη με την βάση δεδομένων – repository η οποία αναπτύχθηκε σε γλώσσα προγραμματισμού Java.

5.2.2.1 C++

Για την ανάπτυξη της μηχανής μετάφρασης το *Microsoft Visual Studio 6* χρησιμοποιήθηκε, συγκεκριμένα αυτό επειδή είχε πολύ καλή συμβατότητα με την βιβλιοθήκη γράφων που εκμεταλλευτήκαμε και αναφέρεται παρακάτω, την *LEDA 4.1*. Η γλώσσα C++ χρησιμοποιήθηκε αντί για την Java (την οποία χρησιμοποιήσαμε για το γραφικό περιβάλλον), για τον λόγο ότι έχει αποτελεσματικότερη διαχείριση μνήμης λόγω δεικτών (και όχι κάποιον garbage collector όπως η Java) και επεξεργαστικά δεν περιλαμβάνει ένα ακόμα επίπεδο επικοινωνίας με το υλικό, όπως η Java (με την *Java Virtual Machine*), άρα και θεωρήσαμε ότι στην περίπτωση μας θα είναι ταχύτερη για τον ρόλο της μηχανής μετάφρασης.

Για την υλοποίηση των αλγορίθμων μετάφρασης, κάποιες επιπλέον βιβλιοθήκες για την διευκόλυνσή μας χρησιμοποιήθηκαν.

- LEDA- 4.1 :

Η βιβλιοθήκη αυτή παρέχει δημιουργία και διαχείριση (εκτός άλλων) γράφων, και

για την συμβατότητά της με το *Microsoft Visual Studio* ό αναγκαστήκαμε να το χρησιμοποιήσουμε. Παρέχει και υποστηρίζει αρκετούς τύπους γράφων καθώς και αποτελεσματική υλοποίηση αποθήκευσης στατικών γράφων. Παρέχει πολλούς αλγόριθμους διάσχισης γράφων (BFS, DFS, Minimal Spanning Trees, Matching, και πολλούς άλλους), ακόμα και ισομορφισμούς γράφων. Υπήρχαν και άλλες εναλλακτικές λύσεις όσον αφορά την διαχείριση γράφων για την γλώσσα C++, αλλά η συγκεκριμένη βιβλιοθήκη κρίθηκε πως παρέχει την μεγαλύτερη ευελιξία.

- BOOST (Phoenix) :

Η υπο- βιβλιοθήκη Phoenix της BOOST είναι βασισμένη σε *συναρτησιακό προγραμματισμό (functional programming)* και είναι επηρεασμένη από την FC++ (Yannis Smaragdakis και Brian McNamara) και την BLL από τους Jaakko Järvi και Gary Powell. Η Phoenix είναι ένα μίγμα αυτών των δυο προσφέροντας και επιπλέον λειτουργίες. Αυτή η βιβλιοθήκη χρησιμοποιήθηκε για την λειτουργία της γραμματικής ανάλυσης των μακροεντολών.

Για την εγκατάσταση και την ορθή λειτουργία των παραπάνω βιβλιοθηκών, χρειάζονται οι εξής ρυθμίσεις στο *Microsoft Visual Studio*¹ :

- i. Tools → Options – Καρτέλα Directories :

- a. Include files : *C:\ALGORITHMIC SOLUTIONS\LEDA-4.1\INCL*
C:\PHOENIX
- b. Library files : *C:\ALGORITHMIC SOLUTIONS\LEDA-4.1*

¹*Παράδειγμα :*

εγκατάσταση LEDA 4.1 : C:\ALGORITHMIC SOLUTIONS

εγκατάσταση PHOENIX : C:\PHOENIX

- ii. Project → Settings – Καρτέλα Link :

- a. Object/Library modules

libD3.lib libw.lib libl.lib libp.lib libg.lib user32.lib gdi32.lib comdlg32.lib shell32.lib advapi32.lib wsock32.lib odbc32.lib

- b. Project Options

libD3.lib libw.lib libl.lib libp.lib libg.lib user32.lib gdi32.lib comdlg32.lib shell32.lib advapi32.lib wsock32.lib odbc32.lib /nologo /subsystem:console /incremental:yes /pdb:"Debug/oraodbc.pdb" /debug /machine:I386 /out:"Debug/oraodbc.exe" /pdbtype:sept

iii. Project → Settings – Καρτέλα C++ :

a. Preprocessor Definitions

`WINMAIN,LEDA_PREFIX`

b. Project Options

`/nologo /MD /Ob2 /D "WINMAIN" /D "LEDA_PREFIX" /Fo"Debug" /O
/Zm300 /TP /c`

5.2.2.2 Java

Η υλοποίηση του γραφικού εργαλείου έγινε αποκλειστικά σε Java – χρησιμοποιώντας Java Development Kit 1.5.0_06 –. Η επιλογή της συγκεκριμένης γλώσσας για την δημιουργία του γραφικού εργαλείου κρίθηκε δόκιμη, αφού η ίδια η Java παρέχει ένα σημαντικό αριθμό κλάσεων που επιταχύνουν την ανάπτυξη γραφικών (μέσω του πακέτου swing και του πιο παλιού awt). Επίσης παράγει κώδικα μεταφέρσιμο σε οποιαδήποτε αρχιτεκτονική. Ο μεταγλωττιστής της Java δεν παράγει κώδικα για την αρχιτεκτονική ενός συγκεκριμένου συστήματος, αλλά για την εικονική μηχανή της Java (Java Virtual Machine). Μια εφαρμογή λοιπόν αυτής της γλώσσας μπορεί να τρέξει σε οποιαδήποτε πλατφόρμα, χωρίς καμιά μετατροπή. Επομένως, ο κώδικας δημιουργείται αλλά και μεταγλωττίζεται μονάχα μια φορά. Επίσης περιλαμβάνει ή υποστηρίζει όλα τα απαραίτητα APIs για την πραγματοποίηση των λειτουργιών που είναι αναγκαίες για την εφαρμογή μας (JDBC – επικοινωνία με τις απαραίτητες βάσεις – εδώ ORACLE – Oracle XE 10.2.0). Αυτή η ευελιξία μας επιτρέπει να μπορούμε να τρέξουμε την εφαρμογή του γραφικού εργαλείου σε οποιοδήποτε σύστημα με εγκατεστημένο το Java Runtime Environment χωρίς καμία απολύτως ρύθμιση.

NetBeans IDE

Το NetBeans είναι το ενοποιημένο περιβάλλον ανάπτυξης εφαρμογών Java που αναπτύσσεται από την Sun Microsystems. Ξεκίνησε από μια φοιτητική εργασία και αγοράστηκε από την Sun η οποία άνοιξε τον κώδικά του.

Ένα σημαντικό σημείο του NetBeans είναι ότι καθώς βασίζεται σε modules (τμήματα λογισμικού) ανεξάρτητες ομάδες μπορούν να προσθέσουν ή να αφαιρέσουν τμήματα από αυτό το περιβάλλον. Έτσι έχουν δημιουργηθεί αρκετές προσθήκες, που προσφέρουν ευελιξία στην εφαρμογή, με συνέπεια να είναι μια από τις δημοφιλέστερες πλατφόρμες ανάπτυξης εφαρμογών. Περιλαμβάνει το *Matisse*, το οποίο είναι μια «προσθήκη» στο NetBeans που παρέχει δημιουργία GUI μέσα από ένα γραφικό περιβάλλον.

- Για την επικοινωνία με την βάση (ORACLE - Oracle XE 10.2.0), είναι αναγκαία η χρήση μιας βιβλιοθήκης που παρέχει τις απαραίτητες κλάσεις. Χρειάζεται να εισάγουμε στο classpath του project NetBeans που έχουμε την κλάση

oracle.jdbc.driver.OracleDriver (το .jar που την περιέχει ονομάζεται συνήθως classes12.jar).

- Επίσης, για την διαχείριση γράφων στην Java, χρησιμοποιήσαμε την βιβλιοθήκη *TG Graph Layout* (TGGraphLayout.jar).

Για την εισαγωγή των βιβλιοθηκών αυτών στο project μας, απλά χρειάζεται :

- i. Project → Properties → Libraries → Add JAR/Folder :
και να εισάγουμε τα δυο προαναφερθέντα .jar αρχεία για την χρήση τους από το project μας.

5.2.3 Σύστημα Διαχείρισης Βάσεων Δεδομένων (ORACLE)

Επιλέξαμε να χρησιμοποιήσουμε ως ΣΔΒΔ την Oracle, το οποίο μαζί με το SQL Server αποτελούν τα πιο εμπορικά Συστήματα Διαχείρισης Βάσεων Δεδομένων που χρησιμοποιούνται σήμερα.

Πιο συγκεκριμένα, χρησιμοποιήθηκε η Oracle Express Edition 10.2.0, και για τις ανάγκες του συστήματος χρειάστηκε να δημιουργηθούν δυο σχήματα, MOVIES και MOVIES_INFO που περιέχουν τα δεδομένα προς μετάφραση και τις πληροφορίες σχετικά με την μετάφραση, αντίστοιχα. Στην περίπτωση μας δημιουργήθηκαν στο ίδιο SID, αλλά αυτό δεν είναι αναγκαίο.

5.2.3.1 TOAD (Quest Software) 8.6.0.38

Για την διαχείριση του συστήματος βάσεων δεδομένων μας, χρησιμοποιήσαμε το εργαλείο TOAD από την Quest Software. Το TOAD σχεδιάστηκε αρχικά από έναν μαθητευόμενο Oracle developer, Jim McDaniel (γνωστός ως "TOADman"), προκειμένου να κατασταθεί η ζωή του ευκολότερη. Το αποκάλεσε "Tool for Oracle Application Developers", TOAD για ευκολία. Παραμένει το επιτυχέστερο εργαλείο διαχείρισης βάσεων δεδομένων Oracle στην κοινότητα των DBAs, παρά τον ανταγωνισμό από τα ίδια τα εργαλεία που προσφέρει η Oracle, και άλλων προμηθευτών. Πρόσφατα, το TOAD έχει επεκταθεί σε άλλες πλατφόρμες βάσεων δεδομένων όπως ο Microsoft SQL Server, DB2, και MySQL..

5.2.3.2 Επιλογή τρόπου επικοινωνίας με τη Βάση Δεδομένων

Για την επικοινωνία με το Σύστημα Διαχείρισης Βάσεων Δεδομένων (εδώ Oracle) χρησιμοποιήθηκαν δυο διαφορετικοί τύποι πρωτοκόλλου. Για την επικοινωνία της βάσης δεδομένων από την μηχανή μετάφρασης, χρησιμοποιήθηκε το πρωτόκολλο ODBC, ενώ για

την διαχείριση της βάσης από την γραφική διεπαφή (το GUI που κατασκευάστηκε σε Java) το πρωτόκολλο JDBC (Java DataBase Connectivity – type 4 – Native Protocol Driver).

- Το JDBC είναι ιδιαίτερα απλό, ευέλικτο και αποτελείται από κλάσεις και διαπροσωπίες (interfaces) που επιτρέπουν την επεξεργασία των δεδομένων μιας οποιασδήποτε Βάσης Δεδομένων με τη χρήση ενσωματωμένης SQL, καθώς επίσης και στην επισκόπηση των μετα-δεδομένων της. Για να μπορέσει μια συγκεκριμένη Βάση Δεδομένων να επικοινωνήσει μέσω JDBC θα πρέπει να υλοποιηθούν οι κατάλληλοι οδηγοί (drivers), κλάσεις δηλαδή, που υλοποιούν τα interfaces και τις αφηρημένες κλάσεις που δηλώνονται από το JDBC API. Αυτό επιτρέπει την πραγματοποίηση μιας άμεσης κλήσης από το μηχανήμα-πελάτη (client machine) στον εξυπηρετητή του ΣΔΒΔ και είναι μια πρακτική λύση για πρόσβαση μέσω κάποιου δικτύου ή ακόμα και μέσω του Internet, κάτι που συμβαίνει στην πλειοψηφία των περιπτώσεων χρησιμοποίησης του συστήματος. Παραπάνω αναφέρθηκαν οι κλάσεις και οι βιβλιοθήκες που χρησιμοποιήθηκαν.
- Το ODBC έχει λιγότερη ευελιξία, αφού χρειάζονται περισσότερες ρυθμίσεις για την ορθή λειτουργία του, όμως είναι μονόδρομος στην λειτουργία της μηχανής μετάφρασης, μιας και λόγω της χρήσης της γλώσσας C++, και δη με τη χρήση του *Microsoft Visual Studio 6**.

* Για την χρήση του ODBC κρίθηκε καλύτερη λύση οι οδηγοί που παρέχει η ίδια η Oracle για ODBC, παρά κάποιου 3^{ου} κατασκευαστή, όπως είναι η Microsoft με την έννοια ότι θα είναι πιο αποδοτικοί, αναγκαίο για την γρήγορη λειτουργία του συστήματος. Οι ρυθμίσεις για την ορθή λειτουργία του ODBC για την χρήση του από την μηχανή μετάφρασης περιλαμβάνουν τα εξής βήματα :

- i. Εγκατάσταση του Oracle 10g Express Edition Client στο σύστημα που θα χρησιμοποιεί την μηχανή μετάφρασης
- ii. Μετά την εγκατάσταση, θα πρέπει στο PATH του συστήματος να υπάρχει ο κατάλογος <ORACLE_HOME>\BIN
- iii. Θα πρέπει το αρχείο TNSNAMES.ORA να υπάρχει στον κατάλογο <ORACLE_HOME>\NETWORK\ADMIN με σωστές παραμέτρους, όπως
XE =
(DESCRIPTION =
(ADDRESS = (PROTOCOL = TCP)(HOST = <DB-Host>)(PORT = 1521))
(CONNECT_DATA =
(SERVER = DEDICATED)
(SERVICE_NAME = XE)
)
)

- iv. Να δημιουργηθεί μια μεταβλητή συστήματος TNS_ADMIN η οποία να δείχνει στον κατάλογο όπου βρίσκεται το TNSNAMES.ORA,
<ORACLE_HOME>\NETWORK\ADMIN
- v. Στα DataSources Management (ODBC) – στα Windows XP Control Panel → Administrative Tools, χρειάζεται στην καρτέλα User DSN να προστεθεί μια εγγραφή με οδηγό Oracle in XE Client (SQORA32.dll) και ρυθμίσεις ότι υπαγορεύει το TNSNAMES.ORA.

6

Έλεγχος

Σε αυτό το κεφάλαιο θα ακολουθήσει η αξιολόγηση του συστήματος, ο τρόπος χρήσης της μηχανής μετάφρασης και του γραφικού εργαλείου για την δημιουργία απαντήσεων, καθώς και θα ακολουθήσουν πειραματικά αποτελέσματα σύμφωνα με την βάση που έχουμε και την βάση- repository που έχουμε κατασκευάσει.

6.1 Τρόπος χρήσης

Ακολουθεί ο τρόπος χρήσης του GUI καθώς και της μηχανής μετάφρασης (μια από τις λειτουργίες του GUI είναι η χρήση της μηχανής μετάφρασης για ευκολία από τον χρήστη).

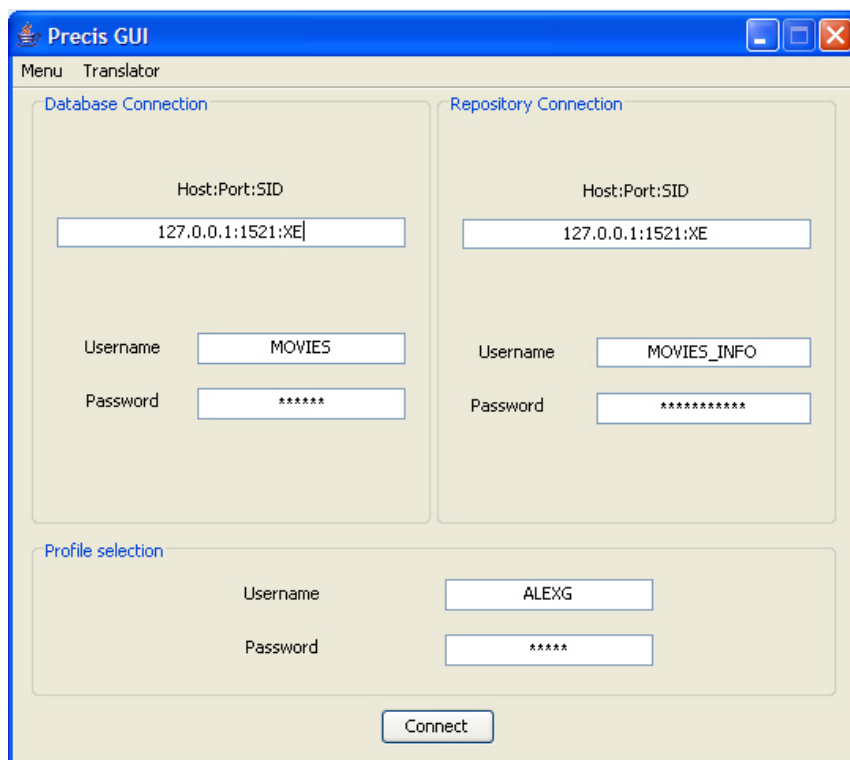
6.1.1 Γραφική διεπαφή χρήστη

6.1.1.1 Αρχική οθόνη

Για να ξεκινήσει η εφαρμογή της διεπαφής χρήστη, αρκεί να τρέξουμε το εκτελέσιμο* `PrecisGUI.jar`, και ευθύς αμέσως θα εμφανιστεί η πρώτη οθόνη, που μας υπαγορεύει να δώσουμε τα απαραίτητα στοιχεία για την μετάφραση, τα οποία είναι

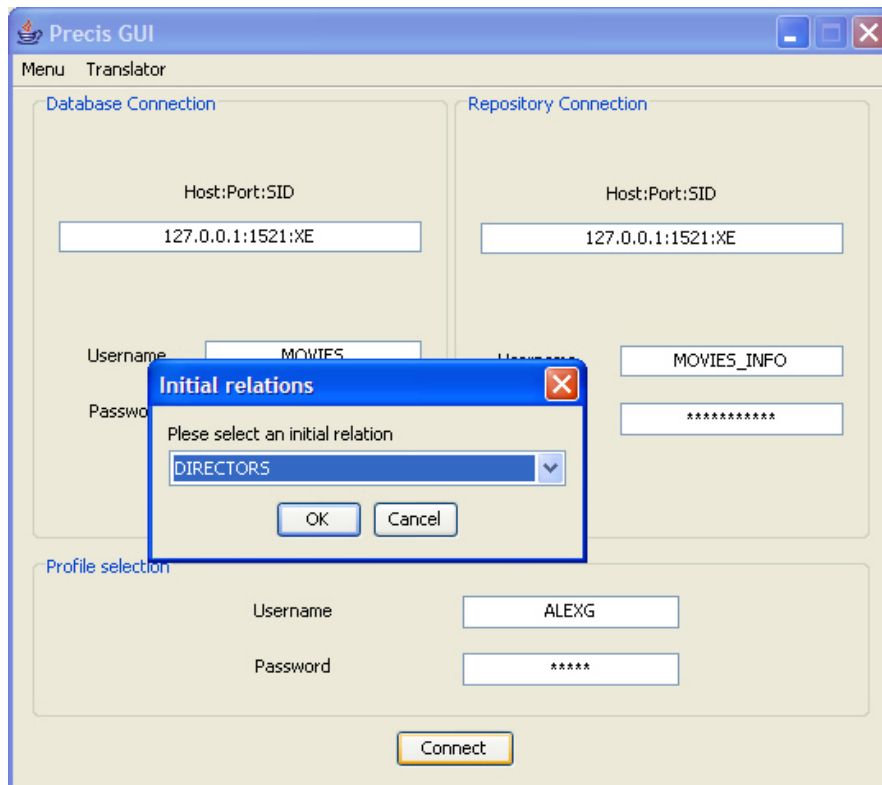
- i. όνομα χρήστη και κωδικό για το σχήμα της βάσης που πρόκειται να μεταφραστεί,
- ii. όνομα χρήστη και κωδικό για την βάση- repository και
- iii. όνομα χρήστη και κωδικό πρόσβασης στο προφίλ του εκάστοτε χρήστη με τις προτιμήσεις του – ετικέτες, μακροεντολές και λοιπά –.

* Εκτελέσιμο .jar αρχείο



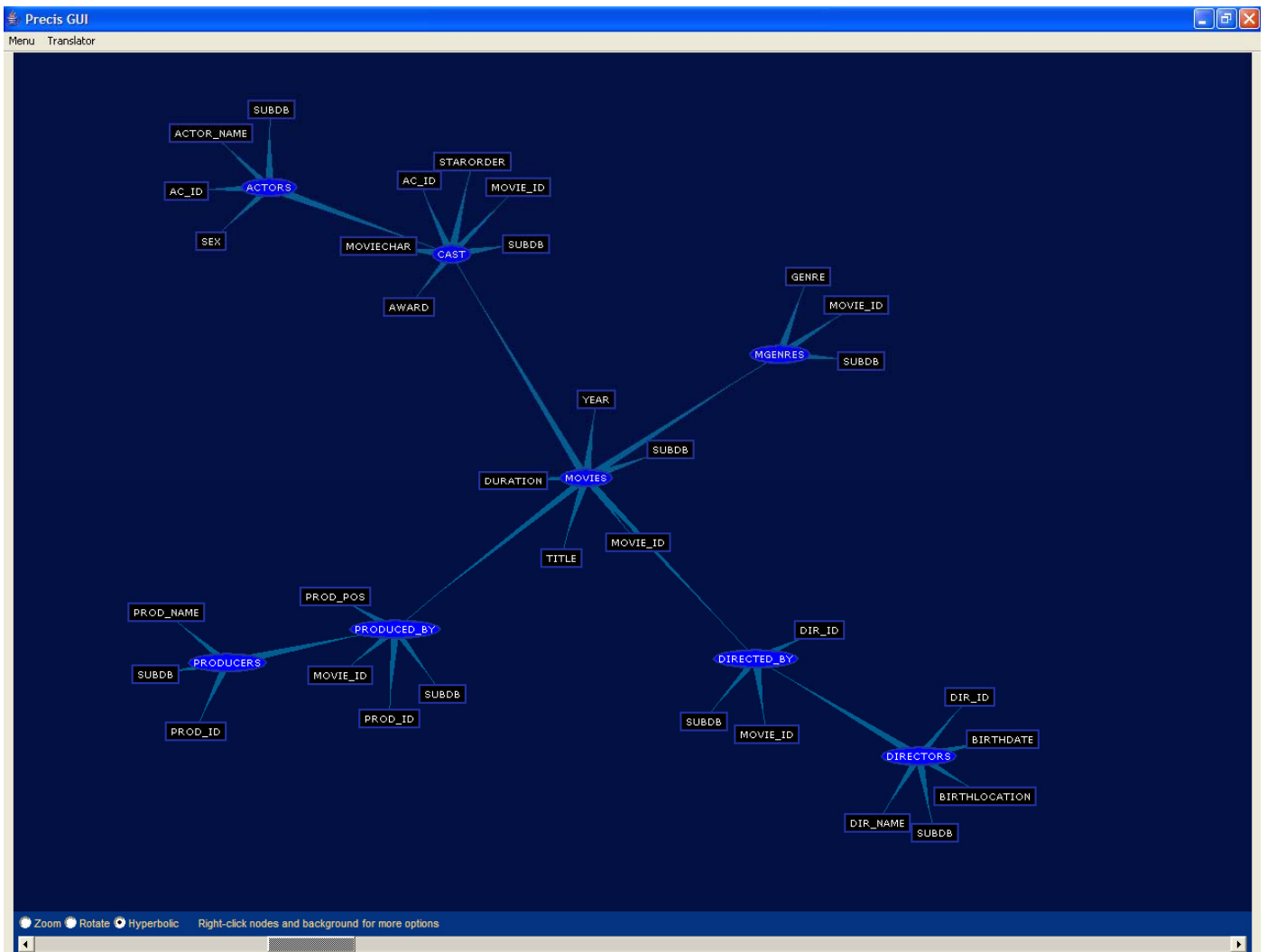
Εικόνα 6.1.1.1.α Οθόνη εισαγωγής στο σύστημα

Αφού εισάγουμε τα απαραίτητα στοιχεία και προσπαθήσουμε να συνδεθούμε, τότε θα μας ερωτήσει για ποια αρχική σχέση επιθυμούμε να αλλάξουμε τα στοιχεία. Υπενθυμίζουμε εδώ ότι μπορεί να έχουμε διαφορετική διάσχιση γράφου αν έχουμε διαφορετική αρχική σχέση, επομένως αλλάζουν και όλα τα στοιχεία, ετικέτες και λοιπά. Επομένως, αν είναι ορισμένες διαφορετικές αρχικές σχέσεις (άρα και διαφορετικοί γράφοι), εδώ επιλέγουμε για ποια αρχική σχέση αλλάζουμε τα στοιχεία του γράφου :



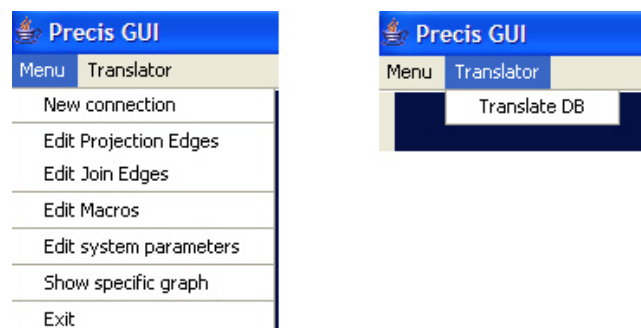
Εικόνα 6.1.1.1.β Οθόνη επιλογής αρχικής σχέσης

Αφού επιλέξουμε και αρχική σχέση, τότε γίνεται η σύνδεση στις δυο βάσεις, την βάση προς μετάφραση και την βάση- repository, και εμφανίζεται μια φόρμα που περιέχει την βάση προς μετάφραση σε γράφο, όπου μπορούμε να επισκοπήσουμε την βάση και να μετακινήσουμε τον γράφο, να κλείσουμε σχέσεις, να κάνουμε ζουμ στον γράφο, απλά για λόγους αισθητικής και κατανόησης του σχήματος της βάσης στον απλό χρήστη.



Εικόνα 6.1.1.1.γ Επισκόπηση βάσης

Στην παρακάτω εικόνα φαίνονται και τα μενού που περιέχει αυτή η αρχική οθόνη, από τα οποία και μπορούμε να κάνουμε ο,τι αλλαγές χρειάζονται για την μετάφραση.

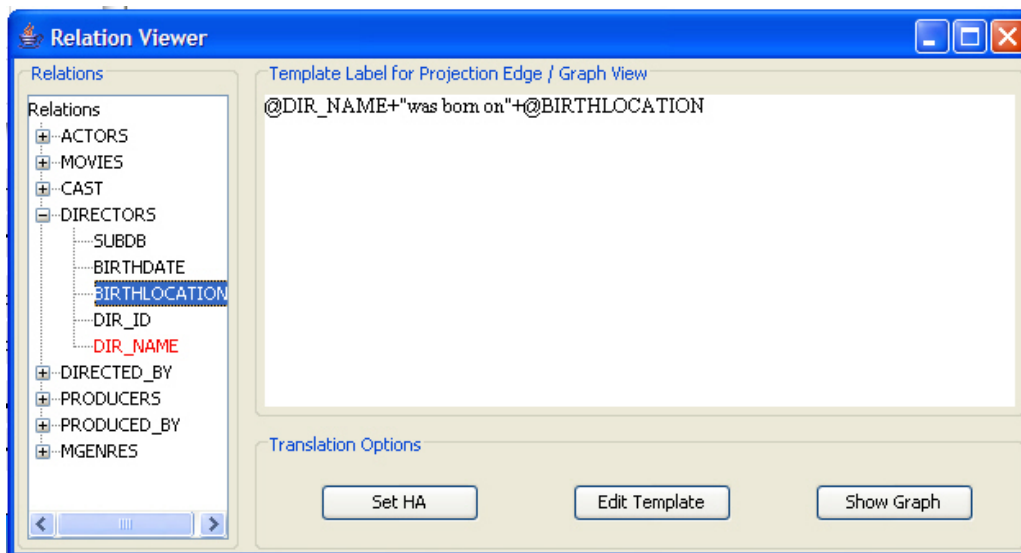


Εικόνα 6.1.1.1.δ Μενού αρχικής οθόνης

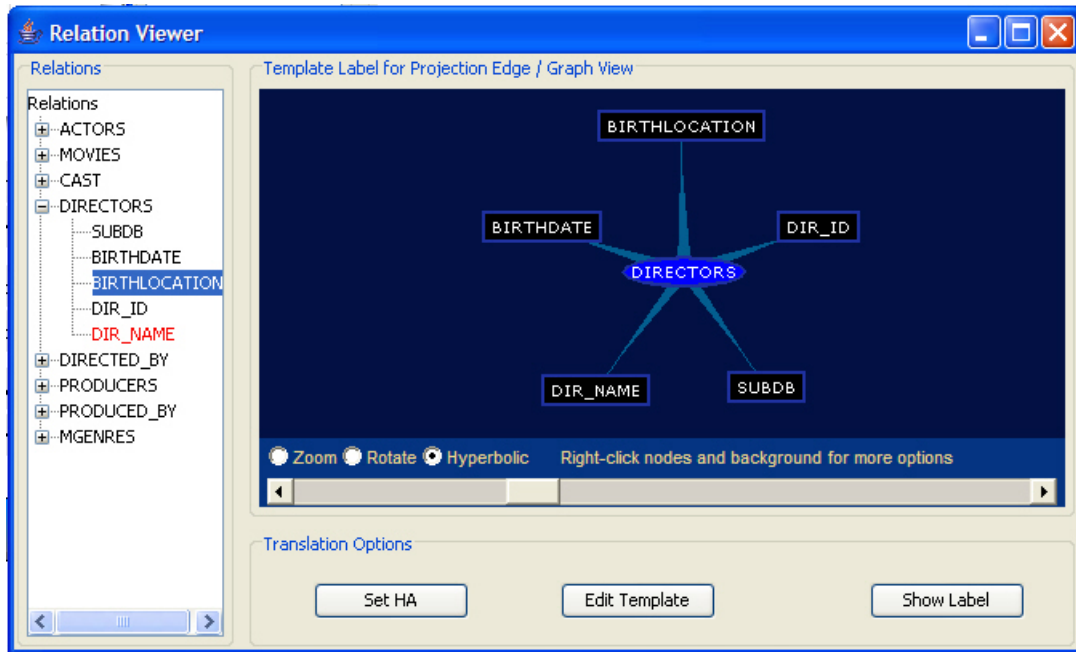
6.1.1.2 Επισκόπηση – αλλαγή ακμών προβολής

Για την επισκόπηση και την αλλαγή ακμών προβολής, αρκεί να επιλέξουμε τη επιλογή από το μενού Edit Projection Edges, όπου και εμφανίζεται μια οθόνη από όπου μπορούμε να δούμε τις σχέσεις και τα πεδία τους σε δεντρική μορφή, τα αρχικά τους πεδία (heading attributes – μαρκαρισμένα κόκκινα) και την ετικέτα που χαρακτηρίζει την ακμή προβολής.

Οι επιλογές μας σε αυτήν την οθόνη είναι η αλλαγή αρχικού πεδίου- HA, όπου και μας ρωτάει αν θέλουμε να απορρίψουμε όλα τα υπόλοιπα στοιχεία που έχουν σχέση με αυτό, μιας και μια αλλαγή αρχικού πεδίου σχέσης έχει αντίκτυπο και σε μακροεντολές και ακμές σύνδεσης εκτός από ακμές προβολής. Επίσης μπορούμε να έχουμε και μια γραφική επισκόπηση της σχέσης που αλλάζουμε κάθε στιγμή.

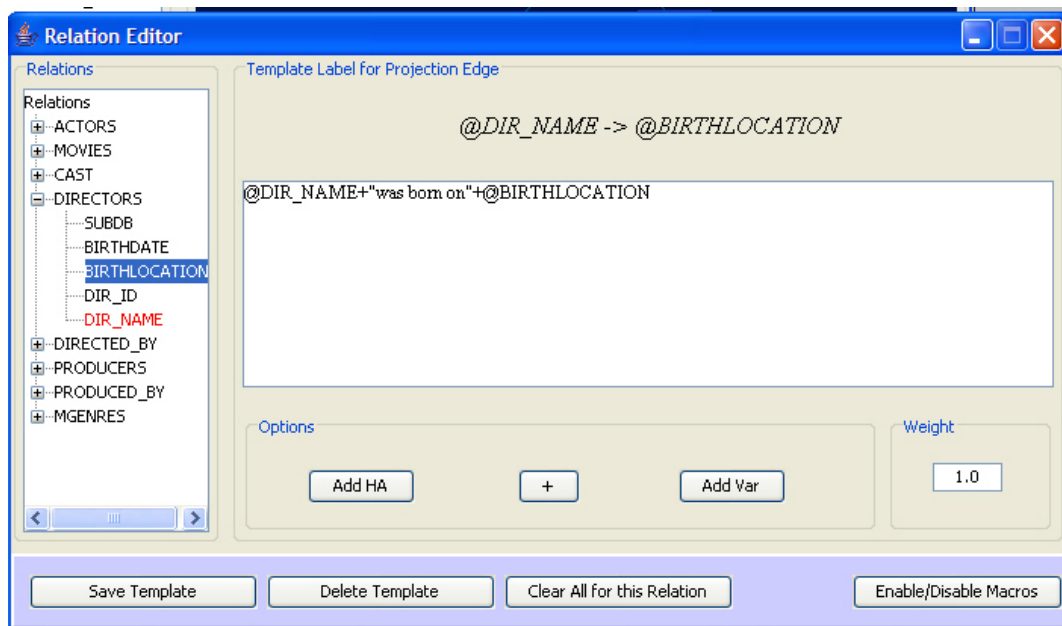


Εικόνα 6.1.1.2.α Επισκόπηση σχέσης

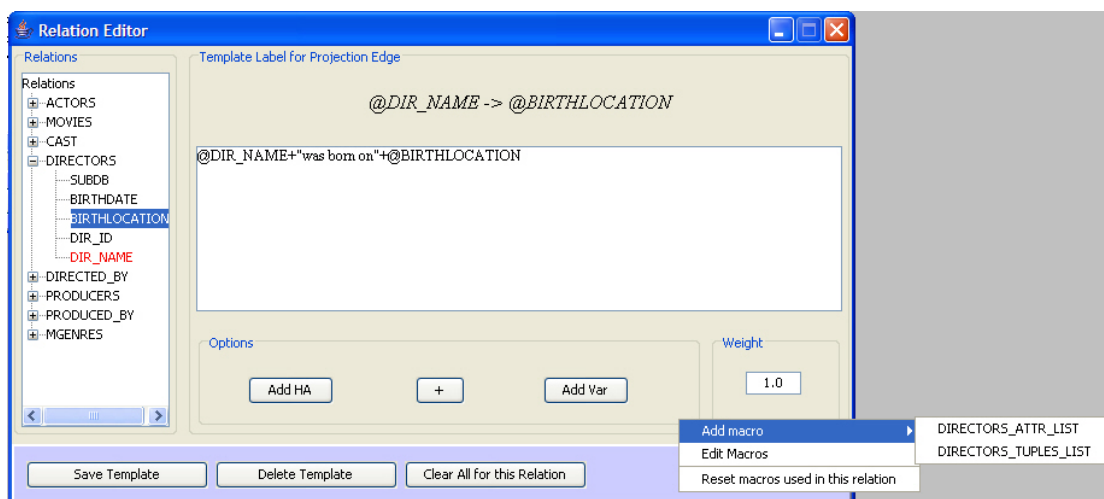


Εικόνα 6.1.1.2.β Επισκόπηση σχέσης σε γραφική μορφή

Αν επιχειρήσουμε να μεταβάλλουμε μια ακμή προβολής, τότε έχουμε τις επιλογές να προσθέσουμε το αρχικό πεδίο (heading attribute – HA) στην ετικέτα της ακμής, να προσθέσουμε κάποιο από τα επιλεγμένα πεδία στο δέντρο, να ορίσουμε το βάρος της ακμής αυτής που αργότερα θα διαβαστεί από την μηχανή μετάφρασης για να αποφασιστεί αν θα μεταφραστεί τελικά, να σώσουμε, να σβήσουμε την τρέχουσα ακμή είτε να σβήσουμε όλες τις ακμές που αναφέρονται στην επιλεγμένη σχέση. Επίσης, μπορούμε είτε να ανοίξουμε την φόρμα μακροεντολών από εδώ, να ενεργοποιήσουμε ή να απενεργοποιήσουμε μακροεντολές για την μετάφραση της σχέσης (μετα- δεδομένα για παράδειγμα). Σε περίπτωση που κάποια μακροεντολή είναι ενεργοποιημένη για την συγκεκριμένη σχέση, τότε βγάζει κατάλληλο μήνυμα, και εμφανίζεται επίσης και ποια μακροεντολή χρησιμοποιείται.



Εικόνα 6.1.1.2.γ Μεταβολή ακμής προβολής

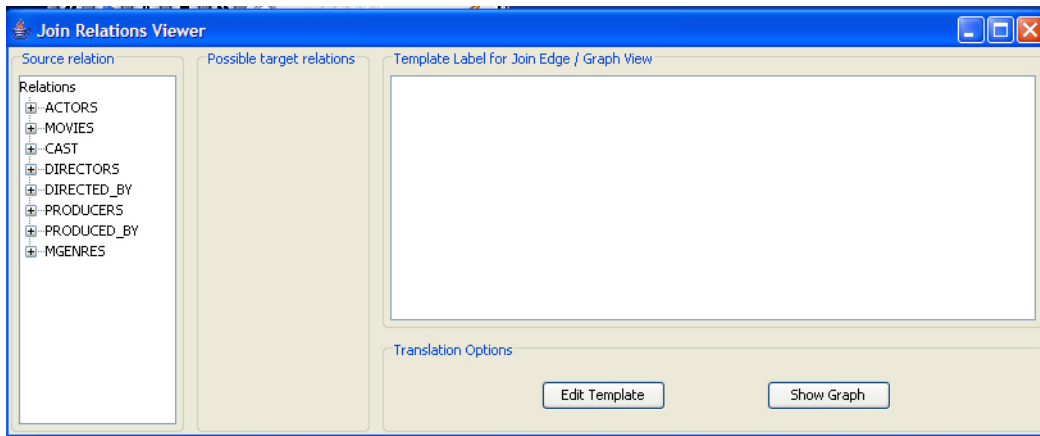


Εικόνα 6.1.1.2.δ Επιλογές μακροεντολών σχέσης

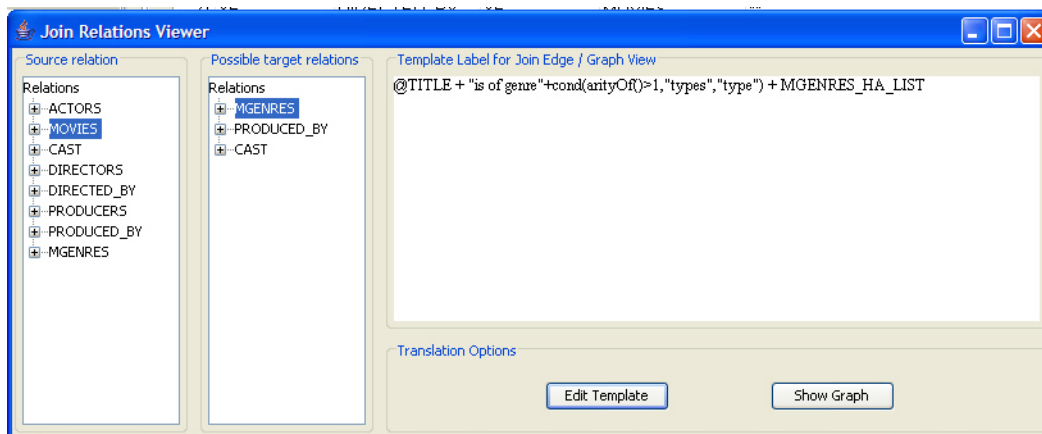
6.1.1.3 Επισκόπηση – αλλαγή ακμών σύνδεσης

Για την επισκόπηση και την αλλαγή ακμών προβολής, αρκεί να επιλέξουμε τη επιλογή από το μενού Edit Join Edges, όπου και εμφανίζεται μια οθόνη από όπου μπορούμε να δούμε τις σχέσεις και τα πεδία τους σε δεντρική μορφή, και αν επιλέξουμε μια από αυτές τις σχέσεις, εμφανίζεται σε δεντρική μορφή οι συνδεδεμένες με αυτήν σχέση, σύμφωνα με την διάσχιση του γράφου που έχουμε επιλέξει στην αρχική οθόνη (σύμφωνα με την αρχική σχέση).

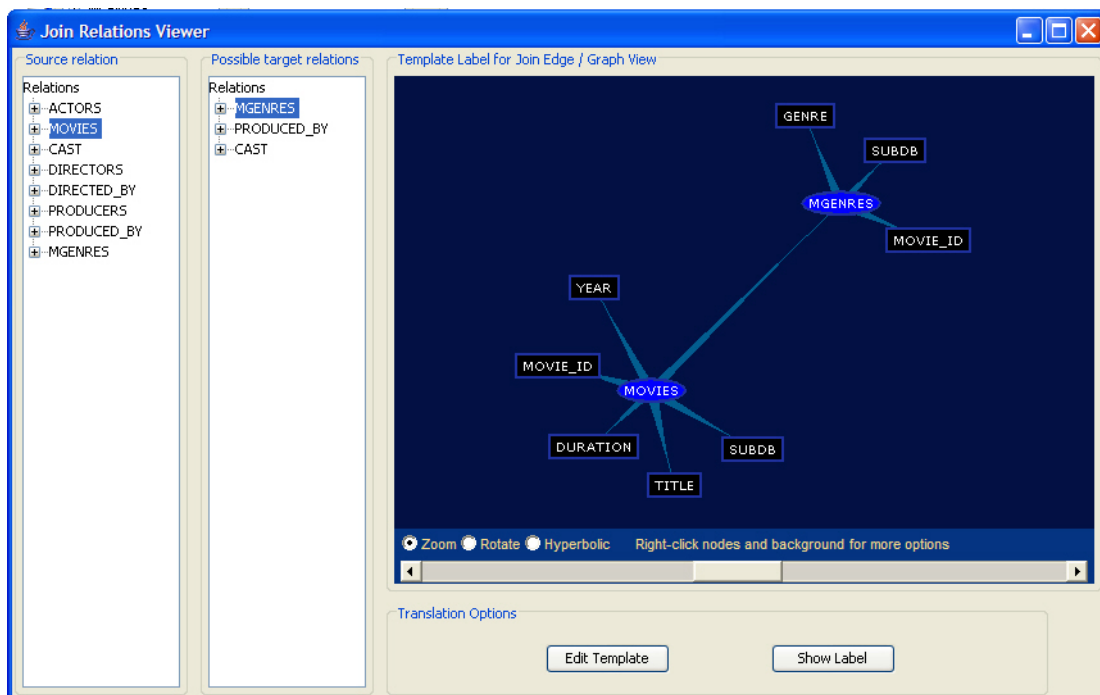
Εδώ οι επιλογές μας είναι είτε να πάμε σε μια φόρμα που αλλάζουμε το περιεχόμενο της ακμής είτε να δούμε τις συνδεδεμένες σχέσεις που έχουμε επιλέξει σε γραφική μορφή.



Εικόνα 6.1.1.3.α Επισκόπηση σχέσεων

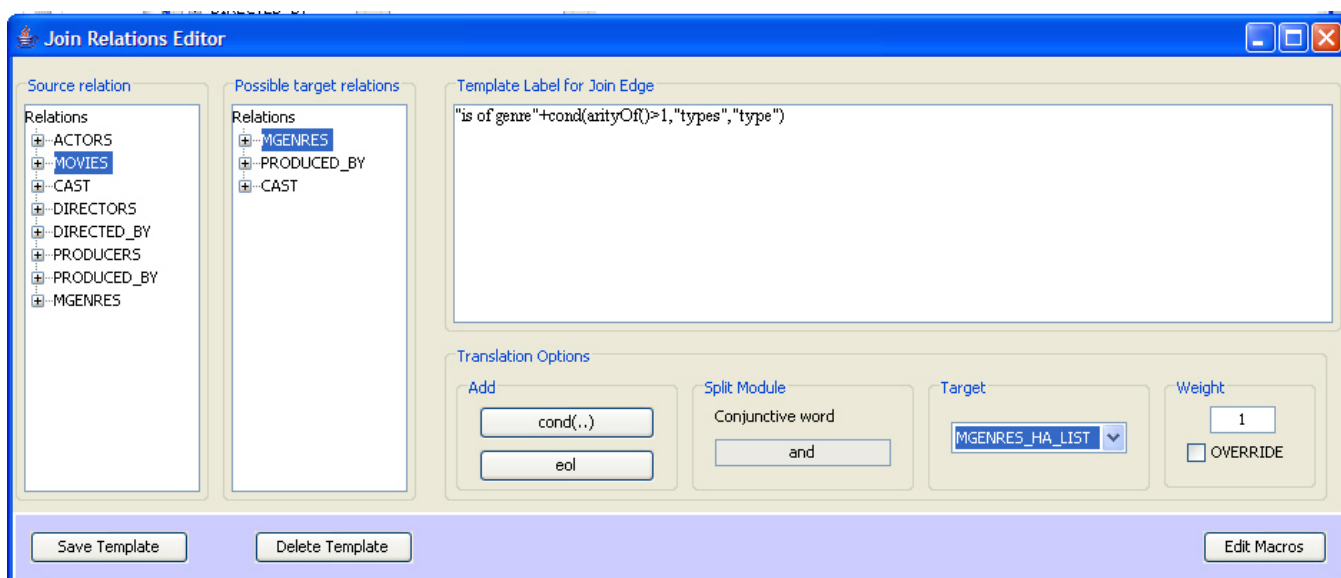


Εικόνα 6.1.1.3.β Επισκόπηση σχέσεων



Εικόνα 6.1.1.3.γ Επισκόπηση σχέσεων σε γραφική μορφή

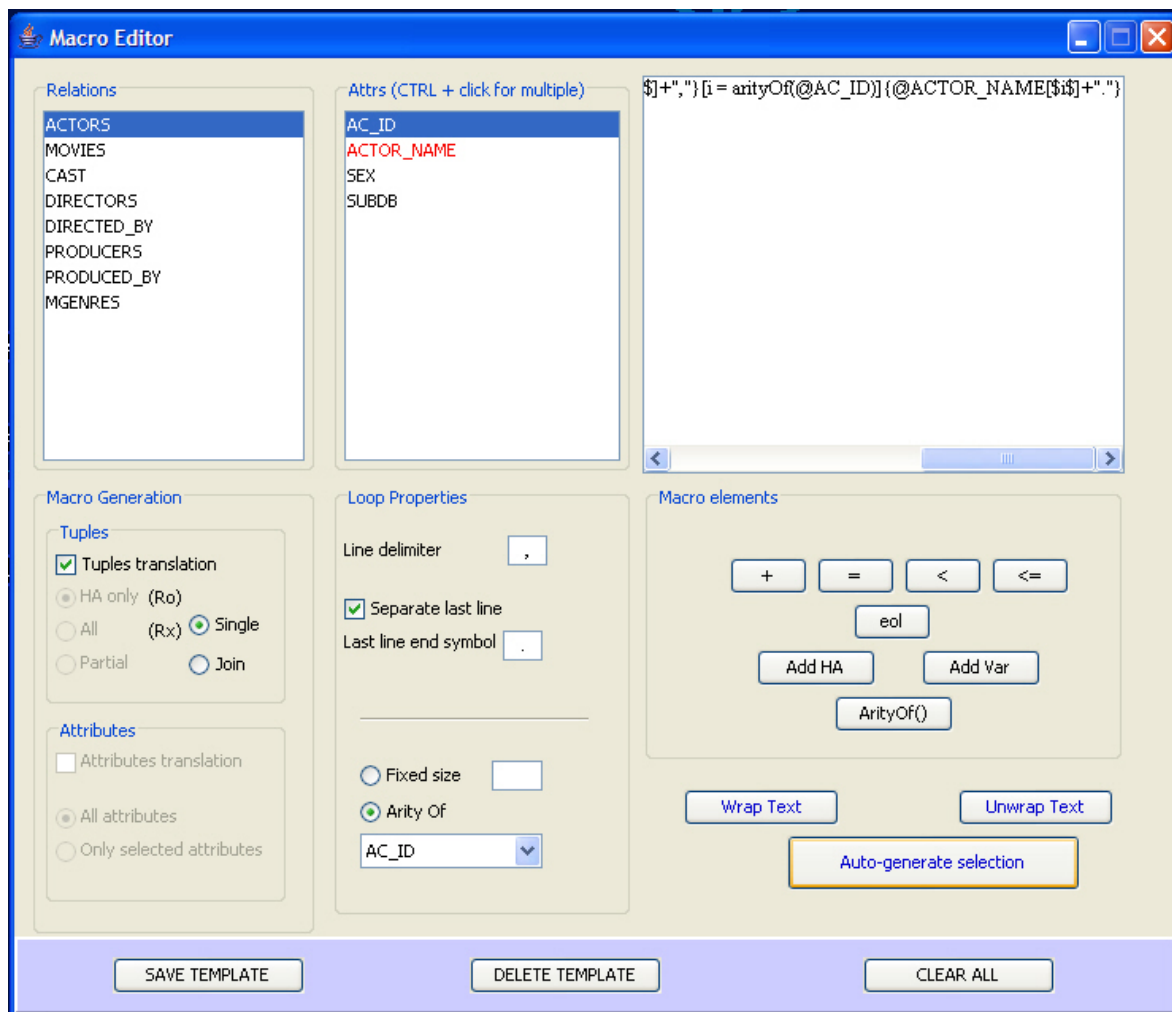
Αν επιχειρήσουμε να αλλάξουμε μια ακμή σύνδεσης σχέσεων, τότε έχουμε την επιλογή της μεταβολής του κειμένου που ακολουθεί το αρχικό πεδίο της σχέσης- πατέρα, και την μακροεντολή που ορίζει την σχέση- παιδί. Επίσης μπορούμε να προσθέσουμε κάποια νέα γραμμή που θα εμφανιστεί στο κείμενο στον χρήστη (σε περίπτωση που θέλουμε παραγράφους – νέες γραμμές στο κείμενο), είτε να ορίσουμε μια νέα συνάρτηση, την `cond(<condition>,<if true>,<if false>)`, όπου ουσιαστικά αν ισχύει το `condition` (μέχρι στιγμής υποστηρίζεται η `arityOf()`) εμφανίζεται το πρώτο όρισμα, αλλιώς το δεύτερο. Αυτό είναι χρήσιμο αν η μακροεντολή που ορίζει την σχέση- παιδί αφορά ενικό ή πληθυντικό, οπότε πρέπει να αλλάξει η πρόταση που προηγείται κατάλληλα. Τέλος, μπορούμε να αλλάξουμε το βάρος της ακμής αυτής αντίστοιχα όπως και στις ακμές προβολής, και να ορίσουμε αν αυτή η σχέση δεν θα μεταφραστεί (`OVERRIDE`), όπως είναι για παράδειγμα οι σχέσεις `PRODUCED_BY`, `DIRECTED_BY`, οι οποίες περιέχουν απλά κάποια πρωτεύοντα κλειδιά και δεν έχουν κάποιο μεταφραστικό ενδιαφέρον. Η συνδετικής λέξη υπάρχει όταν η σχέση- πατέρας που έχει επιλεγεί έχει πολλές σχέσεις- παιδιά, με αποτέλεσμα να μεταφραστεί σε `split module`. Η λέξη δεν μπορεί να αλλάξει από εδώ, είναι ορισμένη στις μεταβλητές συστήματος (παρακάτω ορίζονται).



Εικόνα 6.1.1.3.δ Μεταβολή ακμών σύνδεσης σχέσεων

6.1.1.4 Επισκόπηση – αλλαγή μακροεντολών

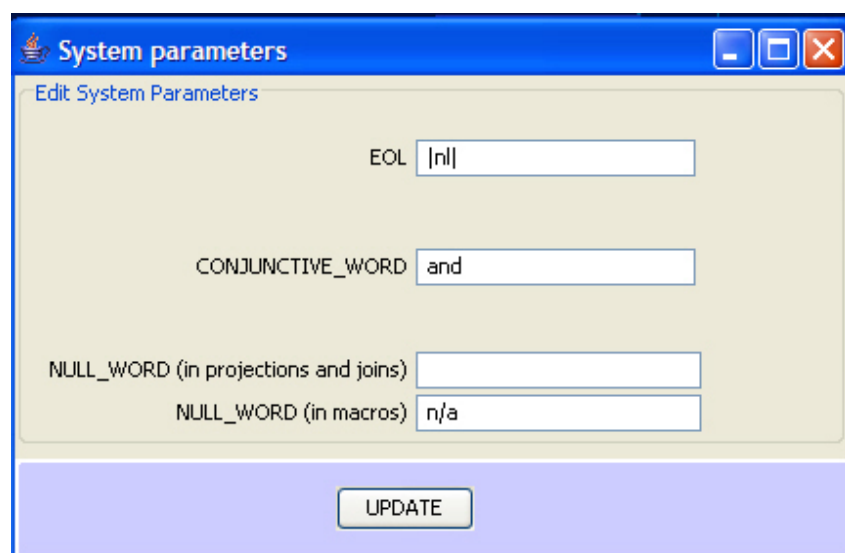
Για την επισκόπηση και την αλλαγή ακμών μακροεντολών, αρκεί να επιλέξουμε τη επιλογή από το μενού Edit Macros, όπου και εμφανίζεται μια οθόνη από όπου μπορούμε να δούμε τις σχέσεις σε ένα παράθυρο, τα πεδία της εκάστοτε επιλεγμένης σχέσης σε ένα άλλο παράθυρο, το κείμενο της κάθε μακροεντολής που ορίζεται από τις τρέχουσες επιλογές του χρήστη (είναι ένα δυναμικό περιβάλλον στο οποίο αν είναι επιλεγμένο για παράδειγμα να μεταβάλλεις μια μακροεντολή που αφορά μετα- δεδομένα της βάσης, τότε εμφανίζεται αυτή στο παράθυρο, αν υπάρχει). Έπειτα υπάρχουν δυο υπο- μενού που ορίζουν αν είναι για μετάφραση μετα- δεδομένων η μακροεντολή (Attributes Translation), ή για σύνδεση μεταξύ σχέσεων (Join Tuples Translation), ή για απλή μετάφραση μιας σχέσης (Single Tuples Translation). Επίσης υπάρχει ένα μενού που αφορά το βρόχο της μακροεντολής, αν θα έχει για παράδειγμα ξεχωριστή γραμμή για την τελευταία πλειάδα που μεταφράζει, και αν θα έχει συγκεκριμένο αριθμό πλειάδων που θα μεταφράσει, όπου πάνω από το όριο αυτό δεν θα μεταφράσει, ή αν θα είναι το μέγεθος του συγκεκριμένου πεδίου (arityOf(@ATTRIBUTE)). Κάτω από το παράθυρο που δείχνει την μακροεντολή, μπορούμε να προσθέσουμε πεδία, τελεστές και ο,τι χρειάζεται μια μακροεντολή για να δημιουργηθεί. Για ευκολία επισκόπησης της μακροεντολής υπάρχει η επιλογή για αναδίπλωση του κειμένου. Επίσης για ευκολία δημιουργίας μακροεντολών από πλευράς χρήστη, έχει δημιουργηθεί μια αυτόματη δημιουργία μακροεντολών, σύμφωνα με τις εκάστοτε επιλογές του χρήστη (αν θα είναι μακροεντολή μετα- δεδομένων και λοιπά). Υπάρχει η επιλογής για αποθήκευση της μακροεντολής στην βάση- repository, η διαγραφή της εκάστοτε επιλεγμένης μακροεντολής από την βάση, είτε ο καθαρισμός του παραθύρου μακροεντολής για δημιουργία νέας.



Εικόνα 6.1.1.4.α Μεταβολή μακροεντολών

6.1.1.5 Επισκόπηση – αλλαγή μεταβλητών συστήματος

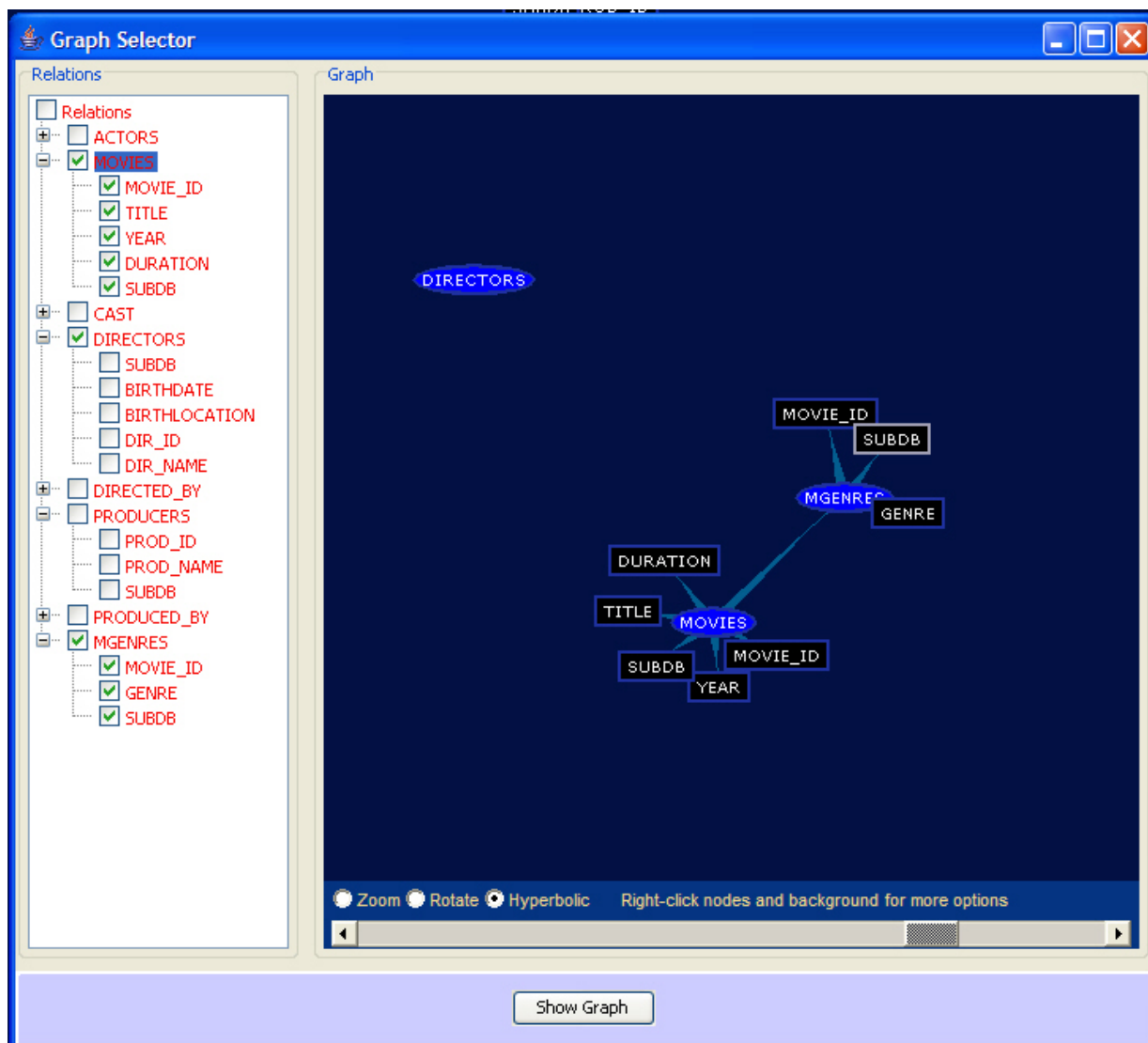
Εδώ υπάρχουν κάποιες μεταβλητές συστήματος (για κάθε προφίλ, φυσικά) που είναι μεν προσκείμενες σε αλλαγή, αλλά δεν ενδείκνυται η συχνή αλλαγή τους, μιας και μπορεί να απασχολούν πολλές ετικέτες, μακροεντολές και ούτω καθ' εξής. Προς το παρόν είναι τέσσερις, μια μεταβλητή που υποδηλώνει πως η μηχανή μετάφρασης θα εισάγει μια νέα γραμμή όταν το βρει, η συνδετική λέξη των σχέσεων που συμμετέχουν σε μια split μετάφραση, το κείμενο που θα εμφανίζεται αν βρεθεί κενή τιμή στην βάση για τις ακμές και το κείμενο που θα εμφανίζεται αν βρεθεί κενή τιμή στην βάση, αλλά για τις μακροεντολές αυτήν την φορά. Όσον αφορά τις ακμές, αν δηλωθεί κενό, τότε θα απορριφθεί ολόκληρη η πρόταση που την περικλείει. Για τις μακροεντολές δεν υπάρχει τέτοια επιλογή, λόγω της φύσης της σύνταξης της μακροεντολής.



Εικόνα 6.1.1.5.α Μεταβολή παραμέτρων συστήματος

6.1.1.6 Επισκόπηση ολικού – μερικού γράφου

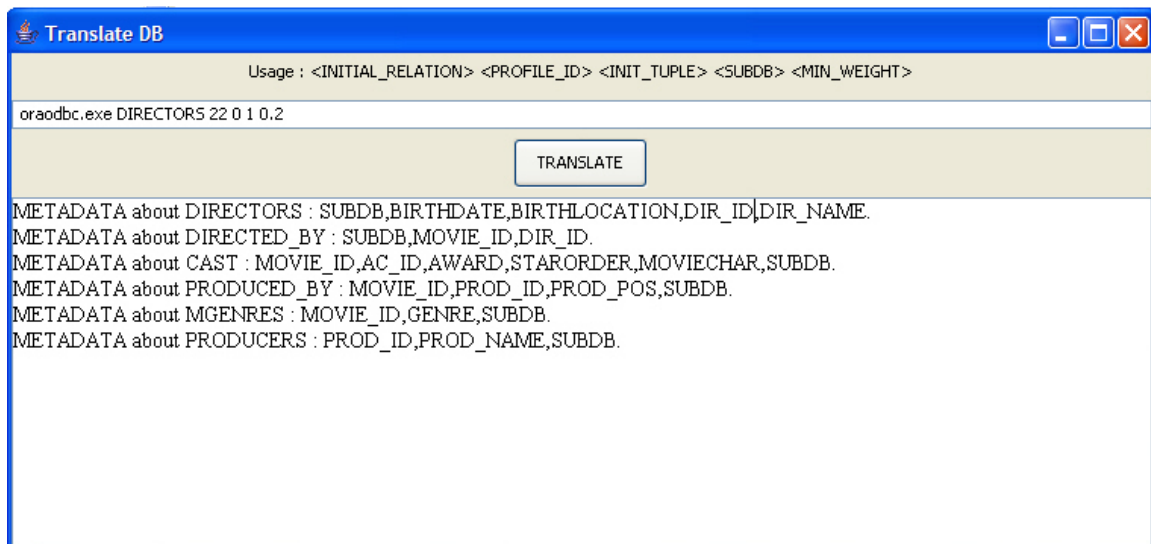
Αυτή η φόρμα μας επιτρέπει επισκόπηση οποιουδήποτε μέρους του γράφου εμείς επιθυμούμε, όποια σχέση θέλουμε με όποια πεδία επιθυμούμε. Γενικώς μια ελευθερία επισκόπησης της βάσης σε μορφή γράφου. Έχουμε επιλογές να ζουμάρουμε στην εικόνα, να αλλάξουμε το σχήμα (σε μορφή υπερβολής), να μετακινήσουμε τις σχέσεις και τα πεδία τους, ή να περιστρέψουμε το σχήμα.



Εικόνα 6.1.1.6.α Επισκόπηση επιλογής γράφου

6.1.1.7 Χρήση μηχανής μετάφρασης

Από το γραφικό εργαλείο είναι δυνατή η χρήση της μηχανής μετάφρασης, αν το εκτελέσιμο αρχείο της μηχανής μετάφρασης βρίσκεται εκεί που βρίσκεται και το εκτελέσιμο αρχείο της γραφικής διεπαφής χρήστη. Από την εικόνα 6.1.1.7.α είναι φανερό πως μπορούμε να εκκινήσουμε κανονικά την μηχανή μετάφρασης με όλα τα ορίσματά της.



Εικόνα 6.1.1.7.α Μηχανή μετάφρασης

6.2 Σενάρια – αποτελέσματα μετάφρασης

Στην ενότητα αυτή παρουσιάζουμε αναλυτικά τα αποτελέσματα μετάφρασης για κάποια σενάρια που θα αναπτυχθούν, επιδεικνύοντας τα περισσότερα χαρακτηριστικά του ζευγαριού συστημάτων, γραφική διεπαφή χρήστη – μηχανή μετάφρασης. Τα παρακάτω παραδείγματα χρησιμοποιούν την διεπαφή ενός χρήστη που υπάρχει δηλωμένος στην βάση και είναι έτοιμος να εισαχθεί στο σύστημα μέσω του GUI και να ξεκινήσει να δημιουργεί ετικέτες, μακροεντολές και ο,τι χρειαστεί για να αρχίσει να βγάζει αποτελέσματα. Όλα σύμφωνα και με τις οδηγίες του παραπάνω κεφαλαίου.

Θα παρουσιάζονται ο εκάστοτε γράφος προς μετάφραση, οι ετικέτες, οι μακροεντολές καθώς και οι λοιπές λεπτομέρειες μετάφρασης που χρησιμοποιεί η μηχανή μετάφρασης για την δημιουργία απάντησης και το αποτέλεσμα που παρουσιάζει η μηχανή μετάφρασης στον χρήστη. Οι λεπτομέρειες δημιουργίας αυτών των ετικετών και παραμέτρων που χρησιμοποιεί η μηχανή μετάφρασης και δημιουργούνται από το γραφικό εργαλείο αναφέρονται αναλυτικά

στο προηγούμενο κεφάλαιο χρήσης των εργαλείων. Θα ξεκινήσουμε από την αρχική σχέση (initial relation) του γράφου όπως έχει δηλωθεί στην βάση- repository από το γραφικό εργαλείο και αργά αργά θα επεκτείνουμε τον γράφο, παρατηρώντας τα αποτελέσματα. Ανάμεσα στα αποτελέσματα θα μεταβάλλουμε τιμές και μεταβλητές συστήματος για να παρατηρήσουμε τις διαφορές που δημιουργούνται στα αποτελέσματα.

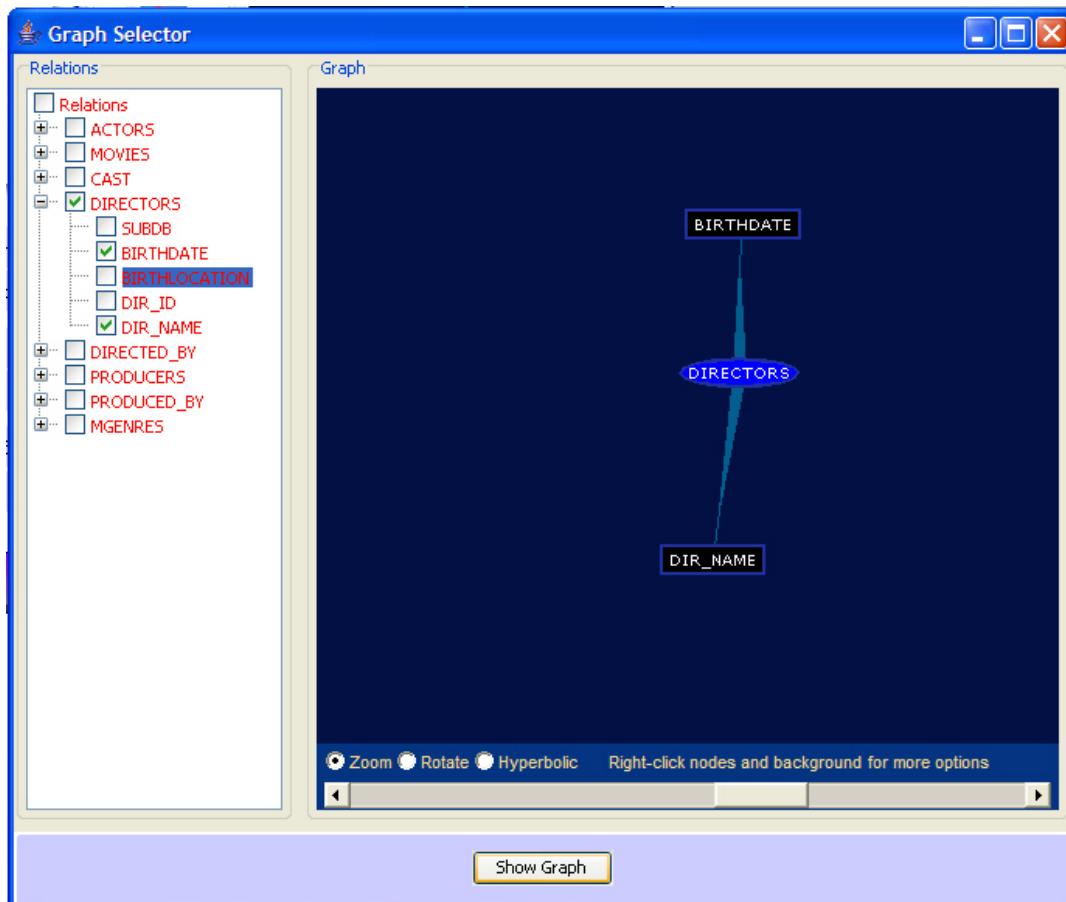
Με την βοήθεια του εργαλείου Graph Selector, εντός του GUI που έχουμε δημιουργήσει, θα επισκοπούμε κάθε προσπάθεια μετάφρασης, ποιους κόμβους και ακμές περιλαμβάνει.

Ξεκινώντας από την αρχική σχέση DIRECTORS, όπως την έχουμε δηλώσει από το γραφικό εργαλείο, καθώς και δυο πεδία (αναγκαστικά το πρωτεύον πεδίο – HA, και άλλο ένα για να παρουσιάσουμε κάποιο αποτέλεσμα), έχουμε την εξής ακολουθία κινήσεων.

Σημείωση : Με h_r συμβολίζουμε το HA (πρωτεύον πεδίο) της σχέσης

6.2.1.1 Μετάφραση σχέσης

Γράφος :



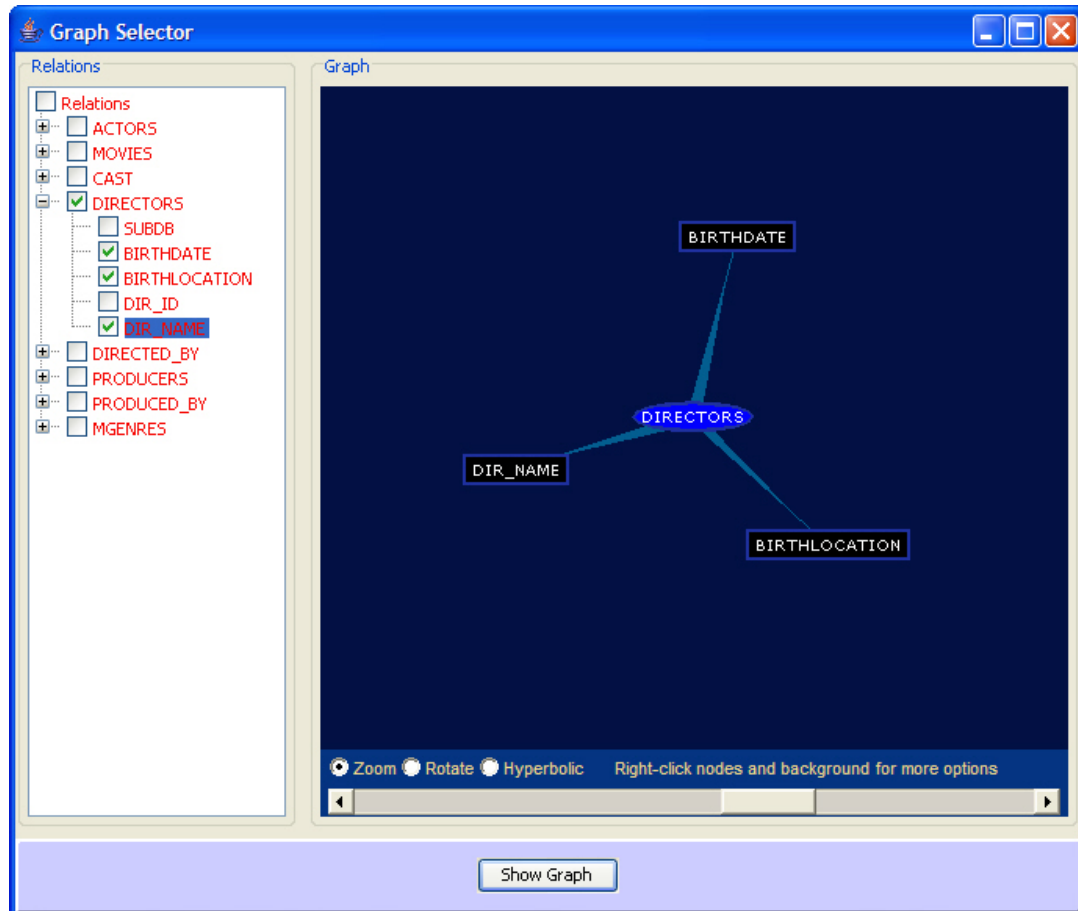
Ετικέτες ακμών προβολής :

$(h_r, BIRTHDATE) = @DIR_NAME + "was born on" + @BIRTHDATE$

Αποτέλεσμα :

Allen, Woody was born on December 1, 1935.

Αν προσθέσουμε και το άλλο πεδίο που περιέχει χρήσιμη πληροφορία,



Ετικέτες ακμών προβολής :

$(h_r, BIRTHDATE) = @DIR_NAME + "was born on" + @BIRTHDATE$

$(h_r, BIRTHLOCATION) = @DIR_NAME + "'s birth location is" + @BIRTHLOCATION$

Αποτέλεσμα :

Allen, Woody 's birth location is Brooklyn, New York, USA. Allen, Woody was born on December 1, 1935.

Αν θελήσουμε να δούμε την λειτουργία της εύρεσης κοινών εκφράσεων/λέξεων, με το οποίο το κείμενο μεταφέρεται πιο κοντά στην φυσική γλώσσα, τότε

Ετικέτες ακμών προβολής :

(h_r, BIRTHDATE) = @DIR_NAME+"was born on"+@BIRTHDATE

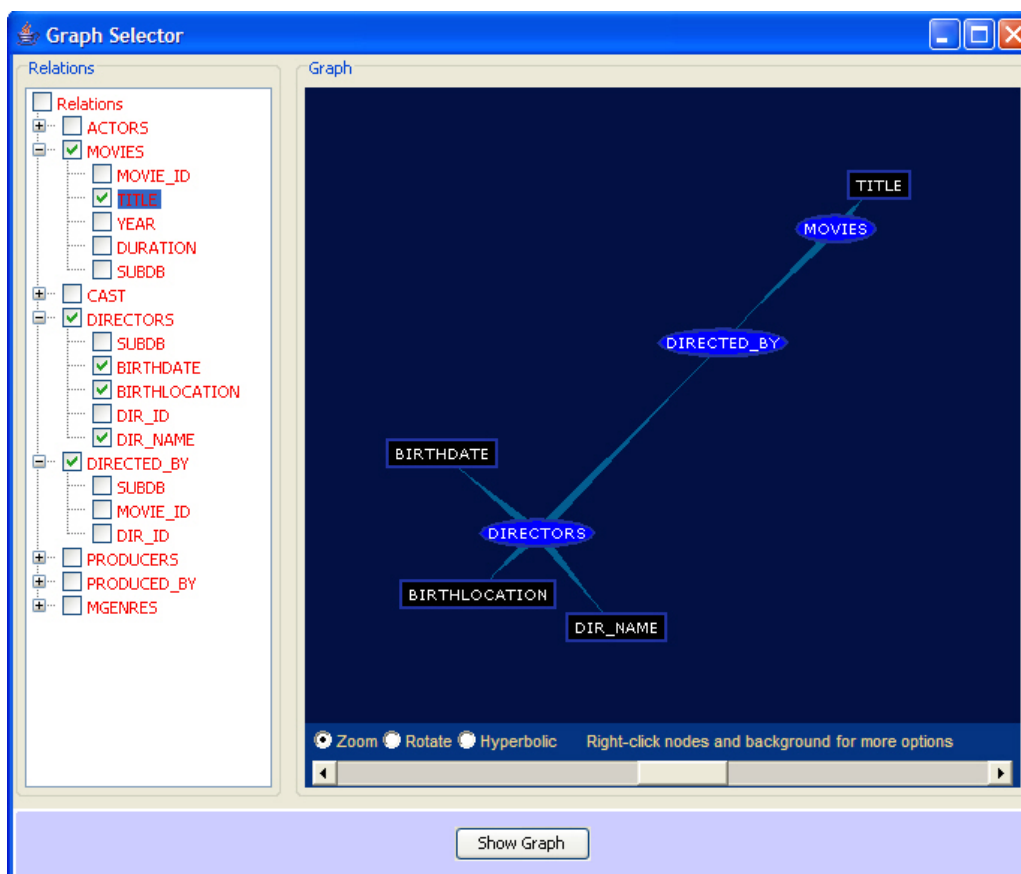
(h_r, BIRTHLOCATION) = @DIR_NAME+" was born in"+@ BIRTHLOCATION

Αποτέλεσμα :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935.

Προχωράμε παρακάτω στις επόμενες συνδεόμενες σχέσεις.

6.2.1.2 *Μετάφραση σύνδεσης σχέσεων, απόρριψη ενδιάμεσης σχέσης, εφαρμογή βαρών*



Η σχέση DIRECTED_BY δεν προσφέρει κανένα ενδιαφέρον στοιχείο που να χρειαζόμαστε να μεταφράσουμε. Προκειμένου να την αποφύγουμε, θα πρέπει να δομηθούν κατάλληλα οι ετικέτες των ακμών σύνδεσης, καθώς και να δηλωθεί στο γραφικό εργαλείο, στην φόρμα που επεξεργαζόμαστε τις ακμές και τις ετικέτες σύνδεσής τους, η αντίστοιχη επιλογή "Override". Για παράδειγμα :

Ετικέτες ακμών προβολής :

```
(hr, BIRTHDATE) = @DIR_NAME+"was born on"+@BIRTHDATE
```

```
(hr, BIRTHLOCATION) = @DIR_NAME+" was born in"+@BIRTHLOCATION
```

Ετικέτες ακμών σύνδεσης :

```
(DIRECTORS, DIRECTED_BY) = @DIR_NAME+"has directed the following  
movies : "+DIRECTED_BY_HA_LIST
```

```
(DIRECTED_BY, MOVIES) = @MOVIE_ID + "" + MOVIES_HA_LIST
```

Μακροεντολές :

```
DEFINE MOVIES_HA_LIST as  
[i <= arityOf(@MOVIE_ID)]  
{@TITLE[$i$]+", "}
```

```
DEFINE DIRECTED_BY_HA_LIST as  
[i <= arityOf(@SUBDB)]  
{""}
```

Με τον παραπάνω τρόπο, το αποτέλεσμα που λαμβάνουμε από την μηχανή μετάφρασης (η οποία παρακάμπτει μεταφραστικά την ενδιάμεση σχέση, αλλά κρατάει τα πρωτεύοντα κλειδιά που χρειάζονται για τις ενώσεις) είναι το εξής :

Αποτέλεσμα :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies : Alice (1990), Annie Hall (1977), Another Woman (1988), Anything Else (2003).

Εάν προσπαθήσουμε να δοκιμάσουμε και την λειτουργία του να προσθέτεις νέες γραμμές μέσα από τις ετικέτες ή τις μακροεντολές, για την εξής μακροεντολή

```
DEFINE MOVIES_HA_LIST as
[i <= arityOf(@MOVIE_ID)]
{"|nl|" + @TITLE[$i$] + ", "}
```

Θα έχουμε αποτέλεσμα :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003).

Εάν στο παραπάνω παράδειγμα ορίσουμε ως το βάρος της ακμής (h, BIRTHLOCATION) είτε ως 0, είτε τρέξουμε την μηχανή μετάφρασης με όρισμα να δέχεται τιμές βάρους μεγαλύτερες από την τιμή που θα επιλέξουμε για την εν λόγω ακμή, τότε το αποτέλεσμα θα είναι :

Allen, Woody was born in Brooklyn, New York, USA. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

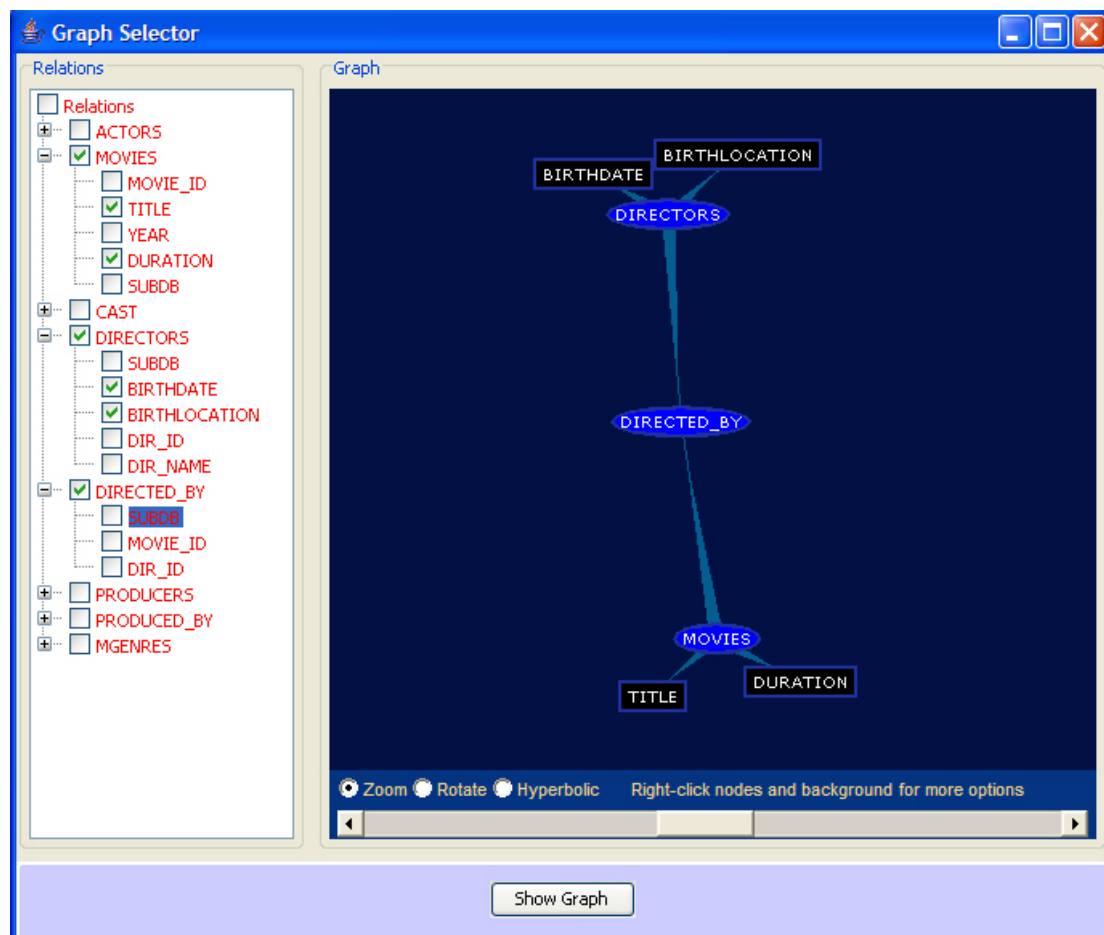
Anything Else (2003).

Εάν επιπλέον σε αυτό ορίσουμε και τιμή για την ακμή (DIRECTORS, DIRECTED_BY) μικρότερη από την επιτρεπόμενη τιμή βάρους στο όρισμα της μηχανής μετάφρασης, το αποτέλεσμα θα είναι :

Allen, Woody was born in Brooklyn, New York, USA.

Οπότε παρατηρούμε την αποκοπή συγκεκριμένων ακμών σύμφωνα με τα βάρη. Στην συγκεκριμένη περίπτωση, αν και έχουμε ορίσει άλλες ακμές, εάν το επιθυμούμε δεν είναι ανάγκη να τις σβήσουμε ή να δημιουργήσουμε ένα άλλο προφίλ, παρά απλά να ορίσουμε τα βάρη που χρειαζόμαστε.

Στις μέχρι τώρα περιπτώσεις όπου υπήρχε τελική σχέση (π.χ. MOVIES), δεν είχαμε δηλώσει κάποια ετικέτα σε ακμή προβολής της τελικής σχέσης αυτής, οπότε και δεν παίρναμε αποτελέσματα για αυτήν (το heading attribute δεν λαμβάνει ετικέτα, επειδή συμμετέχει ενεργά σε όλες τις υπόλοιπες ετικέτες, είτε προβολής, είτε ένωσης). Εάν δηλώσουμε και ετικέτες στις ακμές προβολής της σχέσης αυτής, θα πάρουμε εξής αποτελέσματα :



Ετικέτες ακμών προβολής :

(hr, BIRTHDATE) = @DIR_NAME+"was born on"+@BIRTHDATE

(hr, BIRTHLOCATION) = @DIR_NAME+" was born in"+@ BIRTHLOCATION

(hr, DURATION) = @TITLE+"'s duration is"+@DURATION+"minutes"

Ετικέτες ακμών σύνδεσης :

```
(DIRECTORS, DIRECTED_BY) = @DIR_NAME + "has directed the following  
movies : " + DIRECTED_BY_HA_LIST
```

```
(DIRECTED_BY, MOVIES) = @MOVIE_ID + "" + MOVIES_HA_LIST
```

Μακροεντολές :

```
DEFINE MOVIES_HA_LIST as  
[i <= arityOf(@MOVIE_ID)]  
{ "|nl|" + @TITLE[$i$] + ", " }
```

```
DEFINE DIRECTED_BY_HA_LIST as  
[i <= arityOf(@SUBDB)]  
{ "" }
```

αποτέλεσμα :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Alice (1990) 's duration is 102 minutes. Annie Hall (1977) 's duration is 93 minutes. Another Woman (1988) 's duration is 84 minutes. Anything Else (2003) 's duration is 96 minutes.

6.2.1.3 *Μετάφραση μετα- δεδομένων σχέσεων, χρήση μακροεντολών για μετάφραση απλών σχέσεων*

Εάν, για οποιονδήποτε λόγο, θελήσουμε απλά να έχουμε δεδομένα για το σχήμα της βάσης, λεπτομέρειες για τα δεδομένα των πινάκων, μπορούμε να δηλώσουμε μια μακροεντολή, δομημένη σύμφωνα με την γλώσσα που έχουμε ορίσει και αναγνωρίζει η μηχανή μετάφρασης και να την ενεργοποιήσουμε (μέσα από το γραφικό εργαλείο της διπλωματικής αυτής).

Παράδειγμα : (όλες οι πληροφορίες του αμέσως προηγούμενου +)

Ετικέτες ακμών προβολής :

```
(hr, BIRTHDATE) = @DIR_NAME+"was born on"+@BIRTHDATE
(hr, BIRTHLOCATION) = @DIR_NAME+" was born in"+@ BIRTHLOCATION
(hr, DURATION) = @TITLE+"'s duration is"+@DURATION+"minutes"
```

Ετικέτες ακμών σύνδεσης :

```
DIRECTORS, DIRECTED_BY) = @DIR_NAME + "has directed the following
movies : " + DIRECTED_BY_HA_LIST
DIRECTED_BY, MOVIES) = @MOVIE_ID + "" + MOVIES_HA_LIST
```

Μακροεντολές :

```
DEFINE MOVIES_HA_LIST as
[i <= arityOf(@MOVIE_ID)]
{"|nl|" + @TITLE[$i$] + ", "}

DEFINE DIRECTED_BY_HA_LIST as
[i <= arityOf(@SUBDB)]
{""}

DEFINE MOVIES_ATTR_LIST as
[i < arityOf(@MOVIES)]
{@MOVIES[$i$] + ", "}
[i = arityOf(@MOVIES)]
{@MOVIES[$i$] + "."}
```

Και στο γραφικό εργαλείο, εάν προσπαθήσουμε να δούμε τις πληροφορίες για την συγκεκριμένη σχέση (τις ακμές προβολής), μας ενημερώνει ότι χρησιμοποιείται macro για την μετάφραση της συγκεκριμένης σχέσης και ποια ακριβώς. Το αποτέλεσμα της χρήσης αυτής είναι :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

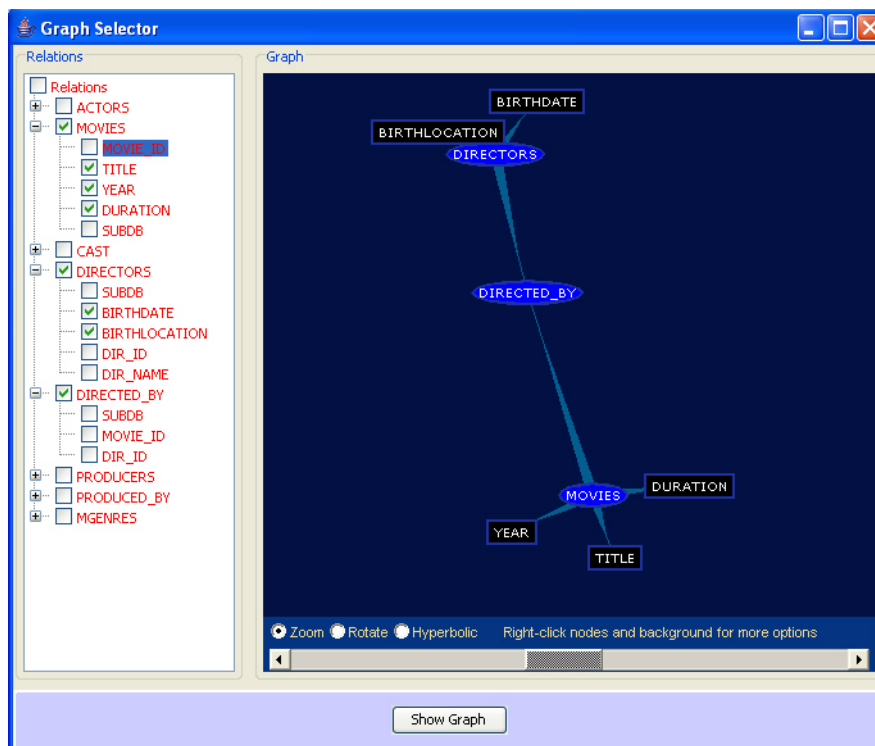
Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). METADATA about MOVIES : MOVIE_ID, TITLE, YEAR, DURATION, SUBDB.

Αν και στο πρωτεύον πεδίο της σχέσης MOVIES εμπεριέχεται η χρονιά που γυρίστηκε η ταινία (αν και δεν χρειάζεται, μιας και στην βάση υπάρχει το πεδίο YEAR το οποίο υποδεικνύει ακριβώς αυτήν την χρονιά), για λόγους επίδειξης της επόμενης δοκιμής, θα ενεργοποιήσουμε και το πεδίο YEAR.

Για την επόμενη δοκιμή, θα δημιουργήσουμε μια μακροεντολή, αλλά αυτήν την φορά όχι για μετα- δεδομένα της βάσης, αλλά για μετάφραση της σχέσης. Αυτό διευκολύνει τον χρήστη εάν η σχέση που επιθυμεί να μεταφράσει έχει πολλά πεδία και θα ήταν πιο χρονοβόρο και κουραστικό να δημιουργεί μια ετικέτα για κάθε πεδίο. Αντίθετα, μπορεί πολύ πιο εύκολα να δημιουργήσει μια μακροεντολή με κείμενο της επιλογής του. Εδώ :



Παράδειγμα : (όλες οι πληροφορίες του αμέσως προηγούμενου, με την διαφορά πως δημιουργούμε και ενεργοποιούμε την εξής μακροεντολή, σε αντίθεση με την αμέσως προηγούμενη που παρουσιάζει τα μετα- δεδομένα της βάσης)

```

DEFINE MOVIES_TUPLES_LIST as
[i <= arityOf(@MOVIE_ID)]
{"Regarding the movie with title "+@TITLE[$i$]+", it was directed on
"+@YEAR[$i$]+" and its duration is "+@DURATION[$i$]+" minutes. "}

```

Αποτέλεσμα μηχανής μετάφρασης :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes.

Στην περίπτωση που ο χρήστης θέλει να έχει κάποια ευελιξία εντός της μακροεντολής, όμως και μετάφραση κάποιων πεδίων μέσω ετικετών, δίνεται και αυτή η επιλογή. Πιο συγκεκριμένα, όσα πεδία δεν περιλαμβάνονται σε κάποια (ενεργοποιημένη) μακροεντολή μετάφρασης σχέσης (όχι μετα- δεδομένων), τότε οι ετικέτες αυτών μεταφράζονται από την μηχανή μετάφρασης κανονικά. Αν, για παράδειγμα μεταβάλλουμε την παραπάνω μακροεντολή και δεν περιλάβουμε ένα πεδίο το οποίο έχει ετικέτα στην βάση δεδομένων (ας πάρουμε για παράδειγμα το πεδίο DURATION, για το οποίο υπάρχει ετικέτα και ας το βγάλουμε από την μακροεντολή), τότε θα έχουμε :

```

DEFINE MOVIES_TUPLES_LIST as
[i <= arityOf(@MOVIE_ID)]
{"Regarding the movie with title "+@TITLE[$i$]+", it was directed on
"+@YEAR[$i$]+"."}

```

Ετικέτα

```
(hr, DURATION) = @TITLE+'s duration is'+@DURATION+'minutes'
```

Τότε το αποτέλεσμα είναι :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Alice (1990) 's duration is 102 minutes. Regarding the movie with title Alice (1990), it was directed on 1990. Annie Hall (1977) 's duration is 93 minutes. Regarding the movie with title Annie Hall (1977), it was directed on 1977. Another Woman (1988) 's duration is 84 minutes. Regarding the movie with title Another Woman (1988), it was directed on 1988. Anything Else (2003)'s duration is 96 minutes. Regarding the movie with title Anything Else (2003),it was directed on 2003.

Σε περίπτωση, φυσικά που δεν υπήρχε καθόλου δηλωμένη ετικέτα για το πεδίο DURATION, η μετάφραση θα ήταν :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

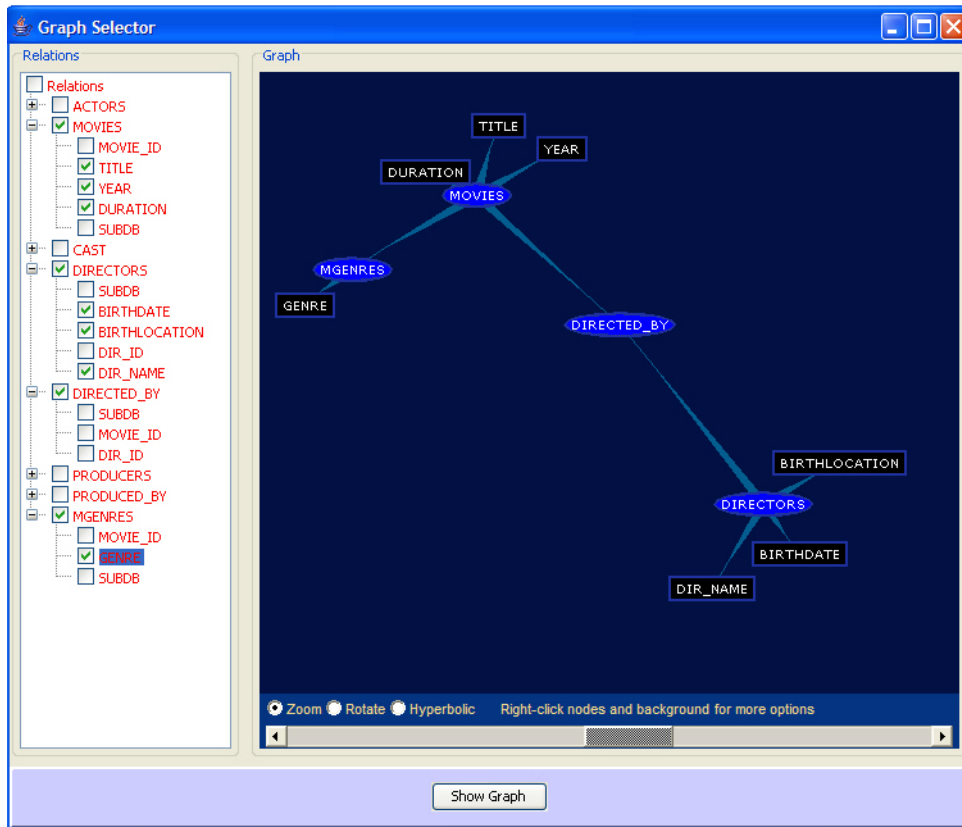
Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990. Regarding the movie with title Annie Hall (1977), it was directed on 1977. Regarding the movie with title Another Woman (1988), it was directed on 1988. Regarding the movie with title Anything Else (2003), it was directed on 2003.

6.2.1.4 *Μετάφραση σύνδεσης σχέσεων, μέσω συνθήκης - cond(..) -*

Ας κρατήσουμε την προηγούμενη μακροεντολή που είχαμε με όλα τα πεδία και ας δηλώσουμε μέσω του γραφικού εργαλείου και άλλη μια σύνδεση με την σχέση GENRES (δηλαδή αν δημιουργήσουμε μια ετικέτα στην ακμή σύνδεσης MOVIES → GENRES ή/και δώσουμε τιμές στις ετικέτες των ακμών προβολής της σχέσης GENRES). Το σχήμα λοιπόν που θα μεταφράσουμε, θα είναι το εξής :



Ανακεφαλαιώνοντας όλες τις ετικέτες, είτε ακμής προβολής είτε ακμής σύνδεσης, και τις εμπλεκόμενες μακροεντολές, θα έχουμε ως μετάφραση του ανωτέρου σχήματος για τις εξής πληροφορίες :

Ετικέτες ακμών προβολής :

(hr, BIRTHDATE) = @DIR_NAME+"was born on"+@BIRTHDATE

(hr, BIRTHLOCATION) = @DIR_NAME+" was born in"+@ BIRTHLOCATION

(hr, DURATION) = @TITLE+"'s duration is"+@DURATION+"minutes"

Ετικέτες ακμών σύνδεσης :

(DIRECTORS, DIRECTED_BY) = @DIR_NAME + "has directed the following movies : " + DIRECTED_BY_HA_LIST

(DIRECTED_BY, MOVIES) = @MOVIE_ID + "" + MOVIES_HA_LIST

(MOVIES, MGENRES) = @TITLE+ "is of genre"+cond(arityOf(>1,"types","type") + MGENRES_HA_LIST

Μακροεντολές :

```
DEFINE MOVIES_HA_LIST as
[i <= arityOf(@MOVIE_ID)]
{"|nl|" + @TITLE[$i$] + ", " }

DEFINE DIRECTED_BY_HA_LIST as
[i <= arityOf(@SUBDB)]
{""}

DEFINE MOVIES_TUPLES_LIST as
[i <= arityOf(@MOVIE_ID)]
{"Regarding the movie with title " + @TITLE[$i$] + ", it was directed on
" + @YEAR[$i$] + " and its duration is " + @DURATION[$i$] + " minutes."}

MGENRES_HA_LIST =
DEFINE MGENRES_HA_LIST as
[i < arityOf(@MOVIE_ID)]
{@GENRE[$i$] + ", " }
[i = arityOf(@MOVIE_ID)]
{@GENRE[$i$] + "." }
```

Σύμφωνα με τις παραπάνω παραμέτρους, λοιπόν, το αποτέλεσμα της μηχανής μετάφρασης είναι :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy,Romance. Regarding the movie with title Another Woman (1988), it was directed on

1988 and its duration is 84 minutes. Another Woman (1988) is of genre type Drama. Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre type Comedy.

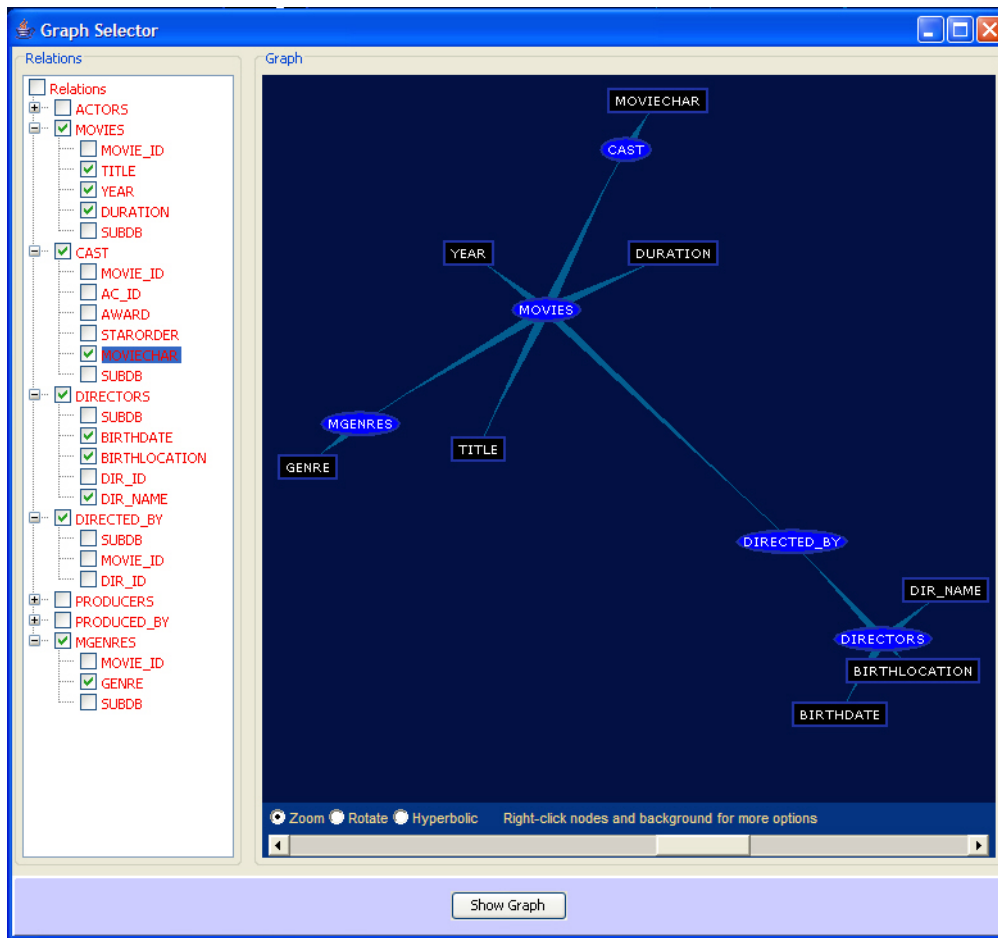
Εδώ παρατηρούμε κάτι καινούργιο, στην ετικέτα ακμής σχέσης που συνδέει την σχέση MOVIES με την νεο- εισερχόμενη σχέση MGENRES, υπάρχει μια συνθήκη- συνάρτηση, με την οποία μπορούμε να ελέγχουμε τις εκφράσεις που υπάρχουν στην ετικέτα. Εδώ, ανάλογα το πλήθος των εγγραφών που επιστρέφει η MGENRES_HA_LIST, βάζουμε στην έκφραση τον ενικό ή τον πληθυντικό της λέξης “type” για την δημιουργία σωστών προτάσεων. Όταν λοιπόν η ταινία υπάγεται σε πάνω από ένα είδος, τότε η έκφραση εκτυπώνει τον πληθυντικό (για παράδειγμα, *Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance.*) ενώ αν η ταινία σχετίζεται με ένα είδος, η έκφραση εκτυπώνει τον ενικό (για παράδειγμα, *Anything Else (2003) is of genre type Comedy.*).

6.2.1.5 Σύνδεση πολλαπλών σχέσεων μέσω unary, περιορισμός αποτελεσμάτων μέσω μακροεντολών

Προχωράμε πιο γοργά και προσθέτουμε στον γράφο προς μετάφραση και την σχέση CAST, για να δούμε και την λειτουργία του split module, δημιουργώντας μια ετικέτα στην ακμή MOVIES → MGENRES. Στην παρούσα περίπτωση, θα μεταφράσουμε μόνο με unary module, και παρακάτω θα ενεργοποιήσουμε την split λειτουργία, για να δούμε και τις διαφορές τους.

Σημείωση : δεν προσθέτουμε στην σχέση MGENRES κάποια ετικέτα μιας και δεν θα προσέφερε τίποτα καινούργιο στην μετάφραση. Η μόνη χρήσιμη πληροφορία που δίνει είναι το πρωτεύον πεδίο της το οποίο και έχει μεταφραστεί από την σύνδεση της MOVIES με την MGENRES.

Ο γράφος είναι :



Τα δεδομένα για την παραπάνω μετάφραση είναι :

Ετικέτες ακμών προβολής :

(hr, BIRTHDATE) = @DIR_NAME+"was born on"+@BIRTHDATE

(hr, BIRTHLOCATION) = @DIR_NAME+" was born in"+@ BIRTHLOCATION

(hr, DURATION) = @TITLE+"'s duration is"+@DURATION+"minutes"

Ετικέτες ακμών σύνδεσης :

(DIRECTORS, DIRECTED_BY) = @DIR_NAME + "has directed the following movies : " + DIRECTED_BY_HA_LIST

(DIRECTED_BY, MOVIES) = @MOVIE_ID + "" + MOVIES_HA_LIST

(MOVIES, MGENRES) = @TITLE + "is of genre"+cond(arityOf(>1,"types","type") + MGENRES_HA_LIST

(MOVIES, CAST) = @TITLE + "had the following cast : " + CAST_HA_LIST

Μακροεντολές :

DEFINE MOVIES_HA_LIST as

[i <= arityOf(@MOVIE_ID)]

{"|nl|" + @TITLE[\$i\$] + ", "}

DEFINE DIRECTED_BY_HA_LIST as

[i <= arityOf(@SUBDB)]

{" "}

DEFINE MOVIES_TUPLES_LIST as

[i <= arityOf(@MOVIE_ID)]

{"Regarding the movie with title " + @TITLE[\$i\$] + ", it was directed on " + @YEAR[\$i\$] + " and its duration is " + @DURATION[\$i\$] + " minutes."}

CAST_HA_LIST =

DEFINE CAST_HA_LIST as

[i < arityOf(@MOVIE_ID)]

{@MOVIECHAR[\$i\$] + ", "}

[i = arityOf(@MOVIE_ID)]

{" and " + @MOVIECHAR[\$i\$] + "."}

MGENRES_HA_LIST =

DEFINE MGENRES_HA_LIST as

[i < arityOf(@MOVIE_ID)]

{@GENRE[\$i\$] + ", "}

[i = arityOf(@MOVIE_ID)]

{@GENRE[\$i\$] + "."}

Το (άκρως αναλυτικό) αποτέλεσμα της μηχανής μετάφρασης είναι :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types

Comedy,Drama,Fantasy,Romance. Alice (1990) had the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter],[Alice at 18

yrs.],[Kimberly],[Monica],[Decorator],[Park Avenue Couples],[Model],[Office Xmas Party Guest],[Dorothy's Maid],[Alice at 12 yrs.],[School Teacher],[Young Dorothy],[Park Avenue Couples],[Kate],[Muse],[Carol],[Circus Equestrian],[Nancy Brill],[Hilda],[Helen],[Circus Aerialist],[Alice's Mother],[Penny],[Dorothy's Xmas Party Guest], and [Nun]. Regarding the

movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes.

Annie Hall (1977) is of genre types Comedy,Romance. Annie Hall (1977) had the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party

Guest],[Janet],[Dorrie],[Allison],[Annie Hall],[Alvy's Classmate],[Grammy

Hall],[Robin],[Alvy's Classmate],[Aunt Tessie],[Actress in Rehearsal],[Annie's Psychiatrist],[Street Stranger],[Street Stranger],[Alvy's Aunt],[School Teacher],[Alvy's Date Outside

Theatre],[Alvy's Mom], and [Waitress at Health Food Restaurant]. Regarding the movie with

title Another Woman (1988), it was directed on 1988 and itsduration is 84 minutes. Another Woman (1988) is of genre type Drama. Another Woman (1988) had the following cast :

[Hope],[Engagement Party Guest],[Waitress],[Cynthia],[Engagement Party

Guest],[Clara],[Young Marion],[Young Claire],[Marion's Mother],[Laura],[Marion

Post],[Young Clara],[Engagement Party Guest],[Little Marion], and [Eleanor Banks].

Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre type Comedy. Anything Else (2003) had the following cast : [NA],[NA], and [NA].

Η απάντηση που λαμβάνουμε ως έξοδο από την μηχανή μετάφρασης μάλλον θα μπορούσε να θεωρηθεί από έναν απλό χρήστη υπερ- αναλυτική. Τα δεδομένα που παρουσιάζονται για τους ηθοποιούς των έργων είναι πολλά, με αποτέλεσμα το μεγαλύτερο μέρος της μετάφρασης να απασχολεί μόνο την απαρίθμηση των ρόλων κάθε ταινίας. Θα μπορούσαμε να κόψουμε εντελώς αυτό το μέρος της μετάφρασης είτε απλά μη δηλώνοντας αυτές τις ετικέτες (άρα και

η μηχανή μετάφρασης δεν θα μετέφραζε τίποτα), είτε αν για κάποια προφίλ επιθυμούσαν αυτήν την πληροφορία να οργάνωναν ανάλογα τα βάρη της μετάφρασης όπως οι ίδιοι επιθυμούν. Για τέτοιου είδους περιπτώσεις, όμως, υπάρχει και μια άλλη λύση που επιτρέπει την οριοθέτηση των αποτελεσμάτων που προέρχονται από μακροεντολές. Αυτό είναι μια λειτουργία που παρέχεται στην δημιουργία μιας μακροεντολής (ή ανανέωσής της), όπου μπορεί να τεθεί ένα άνω όριο μεταφραζόμενων πλειάδων. Πάνω από αυτό το όριο δεν εμφανίζονται στην μετάφραση οι εν λόγω πλειάδες, ούτε λαμβάνουν μέρος στην υπόλοιπη μετάφραση του υπόλοιπου γράφου. Για παράδειγμα, αν από την φόρμα επεξεργασίας μακροεντολών δεν επιλέξουμε την συνάρτηση arityOf(..) για την οριοθέτηση του βρόχου που περιλαμβάνει την μετάφραση των πλειάδων, αλλά αντίθετα, επιλέξουμε έναν σταθερό αριθμό (περισσότερες λεπτομέρειες για την δημιουργία και επεξεργασία των μακροεντολών στο εγχειρίδιο χρήσης), τότε, αν για παράδειγμα θέσουμε έναν αριθμό ορίου ίσο με 4 στην μακροεντολή CAST_HA_LIST, θα έχουμε το εξής αποτέλεσμα (το 4 σημαίνει ότι θα μεταφραστούν ακριβώς 4) :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance. Alice (1990) had the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter],[Alice at 18 yrs.]. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy,Romance. Annie Hall (1977) had the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest],[Janet]. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) is of genre type Drama. Another Woman (1988) had the following cast : [Hope],[Engagement Party Guest],[Waitress],[Cynthia]. Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre type Comedy. Anything Else (2003) had the following cast : [NA],[NA],[NA].

Αν ακόμα μας φαίνονται αρκετά, για μικρότερο αριθμό (για παράδειγμα 2) θα έχουμε :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance. Alice (1990) had the following cast : [Dr. Yang's Assistant],[Alice]. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy,Romance. Annie Hall (1977) had the following cast : [Tony's Date at Nightclub],[Pam]. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) is of genre type Drama. Another Woman (1988) had the following cast : [Hope],[Engagement Party Guest]. Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre type Comedy. Anything Else (2003) had the following cast : [NA],[NA].

6.2.1.6 *Μετάφραση σύνδεσης σχέσεων, με λειτουργία split, εύρεση κοινών εκφράσεων σε πολλαπλά join, σύγκριση*

Αν ενεργοποιήσουμε την λειτουργία του split module, τότε για το ακριβώς παραπάνω παράδειγμα θα έχουμε αποτέλεσμα από την μηχανή μετάφρασης :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types

Comedy, Drama, Fantasy, Romance and has the following cast : [Dr. Yang's Assistant],[Alice]. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy, Romance and has the following cast : [Tony's Date at Nightclub],[Pam]. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) is of genre types Drama and has the following cast : [Hope],[Engagement Party Guest]. Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre types Comedy and has the following cast : [NA],[NA].

Οι διαφορές στις ενώσεις των προτάσεων που δημιουργούνται από τις ετικέτες των ακμών των σχέσεων είναι εμφανείς. Εάν αλλάξουμε την συνδετική λέξη/φράση από το γραφικό εργαλείο για το split module, τότε η μετάφραση παίρνει την εξής μορφή :

(
συνδετική φράση πριν : and
συνδετική φράση μετά : and apart from that
)

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

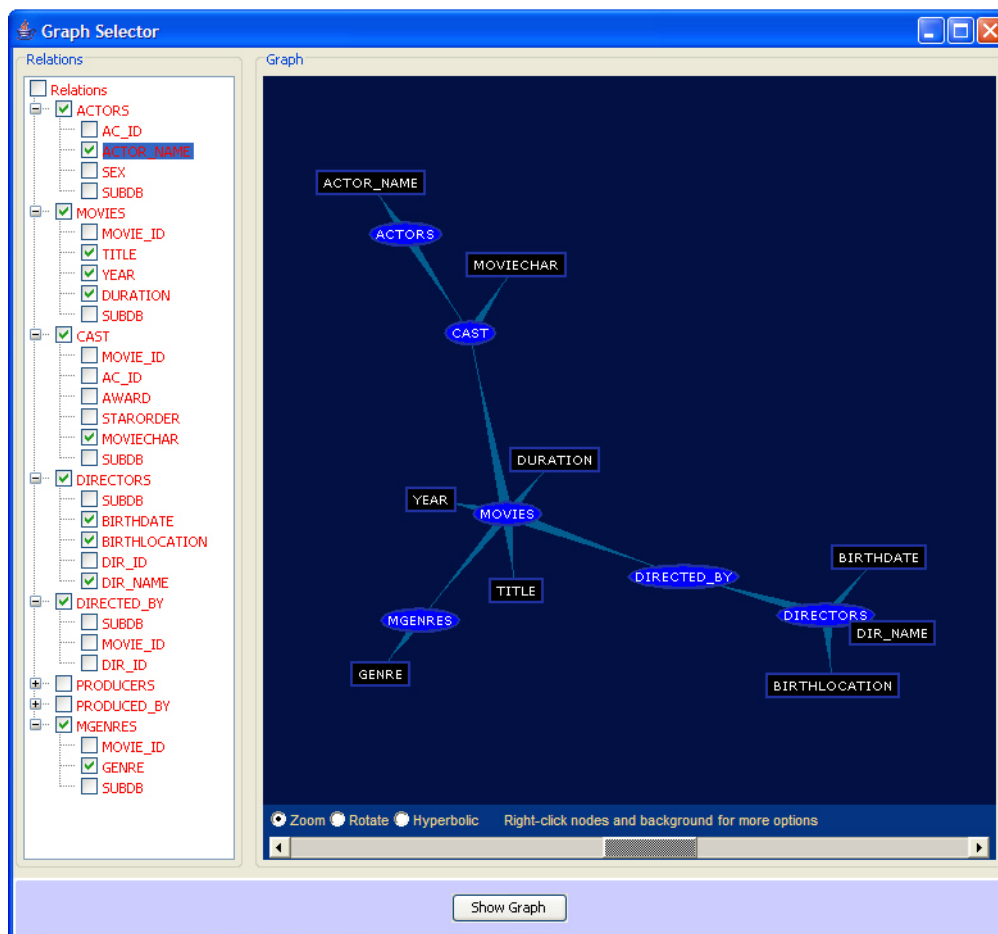
Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy, Drama, Fantasy, Romance and apart from that has the following cast : [Dr. Yang's Assistant],[Alice]. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy, Romance and apart from that has the following cast : [Tony's Date at Nightclub],[Pam]. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) is of genre types Drama and apart from that has the following cast : [Hope],[Engagement Party Guest]. Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre types Comedy and apart from that has the following cast : [NA],[NA].

Επιστρέφουμε στην συνδετική φράση and, και προχωρούμε προσθέτοντας ετικέτες και για τον κόμβο ACTORS στον γράφο μας. Ο γράφος



Ετικέτες ακμών προβολής :

$(h_r, BIRTHDATE) = @DIR_NAME + "was born on" + @BIRTHDATE$

$(h_r, BIRTHLOCATION) = @DIR_NAME + " was born in" + @ BIRTHLOCATION$

$(h_r, DURATION) = @TITLE + "'s duration is" + @DURATION + "minutes"$

Ετικέτες ακμών σύνδεσης :

$(DIRECTORS, DIRECTED_BY) = @DIR_NAME + "has directed the following movies : " + DIRECTED_BY_HA_LIST$

$(DIRECTED_BY, MOVIES) = @MOVIE_ID + "" + MOVIES_HA_LIST$

$(MOVIES, MGENRES) = @TITLE + "is of genre" +$
 $cond(arityOf() > 1, "types", "type") + MGENRES_HA_LIST$


```
(MOVIES, CAST) = @TITLE + "had the following cast : " + CAST_HA_LIST
```

```
(CAST, ACTORS) = @MOVIECHAR + "is in real life" + ACTORS_HA_LIST
```

Μακροεντολές :

```
DEFINE MOVIES_HA_LIST as  
[i <= arityOf(@MOVIE_ID)]  
{ "|n|" + @TITLE[$i$] + ", " }
```

```
DEFINE DIRECTED_BY_HA_LIST as  
[i <= arityOf(@SUBDB)]  
{ "" }
```

```
DEFINE MOVIES_TUPLES_LIST as  
[i <= arityOf(@MOVIE_ID)]  
{ "Regarding the movie with title " + @TITLE[$i$] + ", it was directed on  
" + @YEAR[$i$] + " and its duration is " + @DURATION[$i$] + " minutes." }
```

```
DEFINE CAST_HA_LIST as  
[i < arityOf(@MOVIE_ID)]  
{ @MOVIECHAR[$i$] + ", " }  
[i = arityOf(@MOVIE_ID)]  
{ " and " + @MOVIECHAR[$i$] + "." }
```

```
LIMIT = 2
```

```
DEFINE ACTORS_HA_LIST as  
[i < arityOf(@AC_ID)]  
{ @ACTOR_NAME[$i$] + ", " }  
[i = arityOf(@AC_ID)]  
{ @ACTOR_NAME[$i$] + "." }
```

```
DEFINE MGENRES_HA_LIST as  
[i < arityOf(@MOVIE_ID)]
```

```

{@GENRE[$i$]+", "}
[i = arityOf(@MOVIE_ID)]
{@GENRE[$i$]+". "}

```

Παρατηρούμε ότι στην ετικέτα (CAST, ACTORS), η έκφραση που ενώνει το υποκείμενο (πρωτεύον πεδίο της σχέσης CAST) με το αντικείμενο της πρότασης (η μακροεντολή ACTORS_HA_LIST), αναφέρεται στον ενικό, μιας και ένας χαρακτήρας στο CAST δεν μπορεί να τον υποδύεται πάνω από έναν ηθοποιό. Άρα η μακροεντολή αφορά μόνο μια πλειάδα. Εναλλακτικά, φυσικά, θα μπορούσαμε να χρησιμοποιούμε την συνάρτηση cond(<condition>,<if true>,<if false>), αλλά δεν υπάρχει ανάγκη στην συγκεκριμένη περίπτωση.

Το κείμενο που λαμβάνουμε ως αποτέλεσμα είναι :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

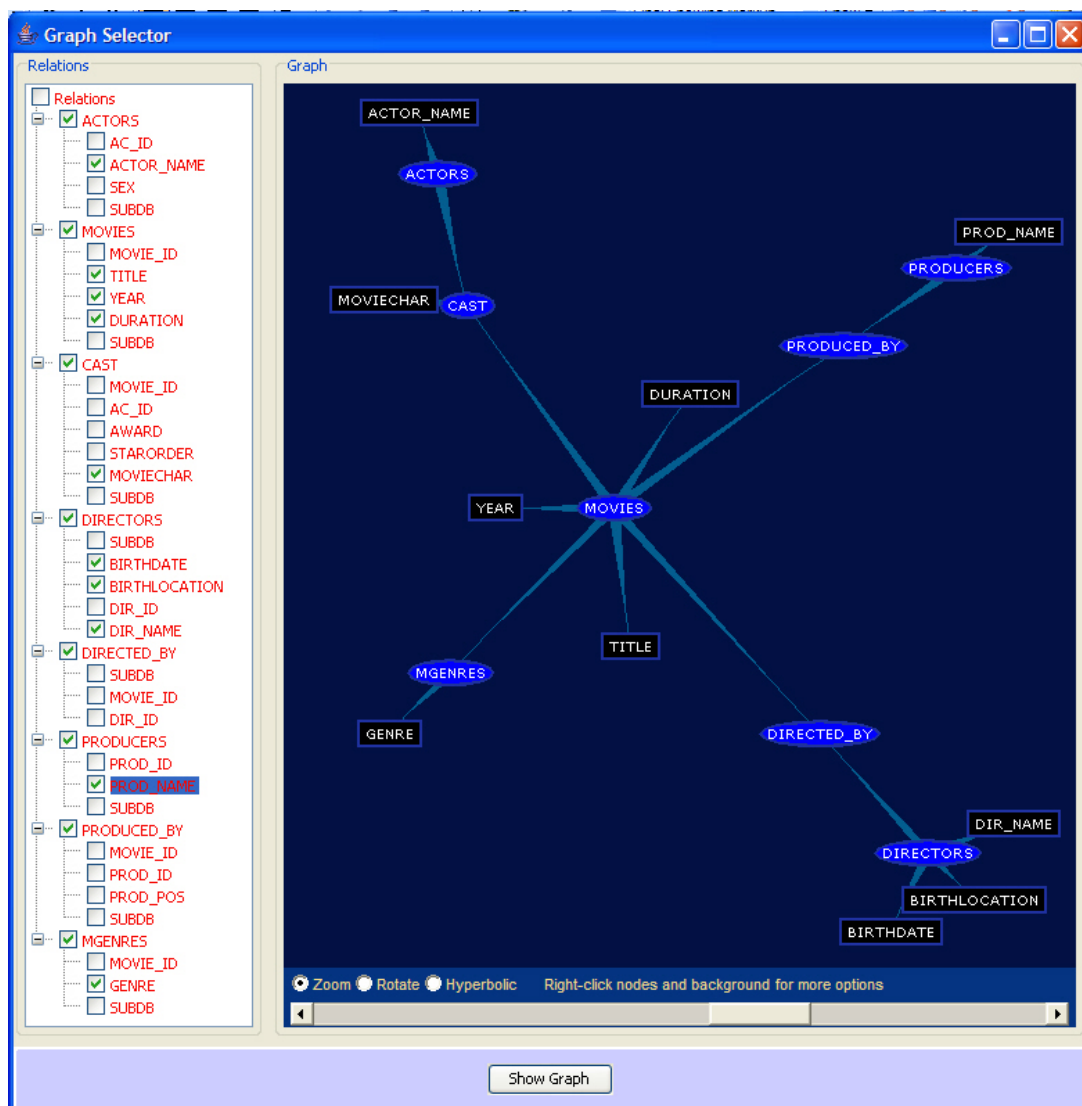
Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance and has the following cast : [Dr. Yang's Assistant],[Alice]. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy,Romance and has the following cast : [Tony's Date at Nightclub],[Pam]. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) is of genre types Drama and has the following cast : [Hope],[Engagement Party Guest]. Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre types Comedy and has the following cast : [NA],[NA]. [Dr.

Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana.

Για να ολοκληρώσουμε τον γράφο, μπορούμε να προσθέσουμε και τις δυο τελευταίες σχέσεις που έχουν απομείνει, την PRODUCED_BY και την PRODUCERS. Δεν έχει νόημα να προσθέσουμε μόνο την PRODUCED_BY, μιας και όπως και η DIRECTED_BY, δεν έχει κάποιο μεταφραστικό ενδιαφέρον, παρά μόνο υπάρχει για την σύνδεση με την PRODUCERS. Αποδεικνύεται πως ούτε η PRODUCERS έχει και πολύ ιδιαίτερο μεταφραστικό ενδιαφέρον, αλλά εδώ εξετάζουμε την λειτουργικότητα της μηχανής μετάφρασης.

Σε αυτήν την περίπτωση θα ενεργοποιήσουμε το unary module για όλες τις σχέσεις, και έπειτα θα εξετάσουμε την λειτουργία με το join module.

Ο εξεταζόμενος γράφος λοιπόν είναι :



Ετικέτες ακμών προβολής :

```
(hr, BIRTHDATE) = @DIR_NAME+"was born on"+@BIRTHDATE
(hr, BIRTHLOCATION) = @DIR_NAME+" was born in"+@ BIRTHLOCATION
(hr, DURATION) = @TITLE+"'s duration is"+@DURATION+"minutes"
```

Ετικέτες ακμών σύνδεσης :

```
(DIRECTORS, DIRECTED_BY) = @DIR_NAME + "has directed the following
movies : " + DIRECTED_BY_HA_LIST
```

```
(DIRECTED_BY, MOVIES) = @MOVIE_ID + "" + MOVIES_HA_LIST
```

```
(MOVIES, MGENRES) = @TITLE + "is of
genre"+cond(arityOf(>1,"types","type") + MGENRES_HA_LIST
```

```
(MOVIES, CAST) = @TITLE + "had the following cast : " + CAST_HA_LIST
```

```
(CAST, ACTORS) = @MOVIECHAR + "is in real life" + ACTORS_HA_LIST
```

```
(MOVIES, PRODUCED_BY) = @TITLE + "was produced by " +
PRODUCED_BY_HA_LIST
```

```
(PRODUCED_BY, PRODUCERS) = @MOVIE_ID + "" + PRODUCERS_HA_LIST
```

Μακροεντολές :

```
DEFINE MOVIES_HA_LIST as
```

```
[i <= arityOf(@MOVIE_ID)]
```

```
{"|nl|"+@TITLE[$i$]+", "}
```

```
DEFINE DIRECTED_BY_HA_LIST as
```

```
[i <= arityOf(@SUBDB)]
```

```
{""}
```

```
DEFINE MOVIES_TUPLES_LIST as
```

```
[i <= arityOf(@MOVIE_ID)]
```

```
{"Regarding the movie with title "+@TITLE[$i$]+", it was directed on
"+@YEAR[$i$]+" and its duration is "+@DURATION[$i$]+" minutes."}
```

```

DEFINE CAST_HA_LIST as
[i < arityOf(@MOVIE_ID)]
{@MOVIECHAR[$i$]+", "} [i = arityOf(@MOVIE_ID)]
{" and "+@MOVIECHAR[$i$]+". "}
LIMIT = 2

```

```

DEFINE ACTORS_HA_LIST as
[i < arityOf(@AC_ID)]
{@ACTOR_NAME[$i$]+", "}
[i = arityOf(@AC_ID)]
{@ACTOR_NAME[$i$]+". "}

```

```

DEFINE PRODUCED_BY_HA_LIST as
[i <= arityOf(@MOVIE_ID)]
{" "}

```

```

DEFINE PRODUCERS_HA_LIST as
[i <= arityOf(@PROD_ID)]
{@PROD_NAME[$i$]+", "}

```

```

DEFINE MGENRES_HA_LIST as
[i < arityOf(@MOVIE_ID)]
{@GENRE[$i$]+", "}
[i = arityOf(@MOVIE_ID)]
{@GENRE[$i$]+". "}

```

Το κείμενο που λαμβάνουμε ως αποτέλεσμα είναι :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice]. Alice (1990) was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy,Romance. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam]. Annie Hall (1977) was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. AnotherWoman (1988) is of genre type Drama. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest]. Another Woman (1988) was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre type Comedy. Anything Else (2003) has the following cast : [NA],[NA]. Anything Else (2003) was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana.

Αν τώρα, για τα ίδια δεδομένα, ενεργοποιήσουμε και χρησιμοποιήσουμε το split module, αντί για το unary παντού, με αποτέλεσμα η μετάφραση των MOVIES → MGENRES, MOVIES → CAST, MOVIES → PRODUCED_BY να συμπυκωθεί και να γίνει σε ένα βήμα, τότε θα έχουμε από την μηχανή μετάφρασης το εξής κείμενο :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast: [Dr. Yang's Assistant],[Alice] and is of genre types Comedy,Drama,Fantasy,Romance and was produced

by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title *Annie Hall* (1977), it was directed on 1977 and its duration is 93 minutes. *Annie Hall* (1977) has the following cast : [Tony's Date at Nightclub],[Pam] and is of genre types Comedy,Romance and was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title *Another Woman* (1988), it was directed on 1988 and its duration is 84 minutes. *Another Woman* (1988) has the following cast : [Hope],[Engagement Party Guest] and is of genre types Drama and was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title *Anything Else* (2003), it was directed on 2003 and its duration is 96 minutes. *Anything Else* (2003) has the following cast : [NA],[NA] and is of genre types Comedy and was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana.

Παρατηρούμε ότι οι τρεις προτάσεις στο unary module,

Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance. *Alice* (1990) has the following cast : [Dr. Yang's Assistant],[Alice]. *Alice* (1990) was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II).

έγιναν μια στην μετάφραση με split module,

Alice (1990) has the following cast: [Dr. Yang's Assistant],[Alice] and is of genre types Comedy,Drama,Fantasy,Romance and was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II).

Με το παραπάνω γίνεται εμφανές ότι το split module παράγει προτάσεις που προσεγγίζουν καλύτερα φυσικής μορφής από ότι το unary. Άλλωστε, αυτή είναι και η διαφορά τους και για αυτό το λόγο δημιουργήθηκε το split module.

Για λόγους πληρότητας ακολουθούν οι μεταφράσεις σε unary και split mode όλου του γράφου, χωρίς τα όρια των δυο (2) πλειάδων που έχουμε ορίσει στο CAST_HA_LIST.

Κείμενο γράφου σε unary mode, χωρίς όρια στις μακροεντολές :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance. Alice (1990) has the following cast :

[Dr. Yang's Assistant],[Alice],[Joe's Daughter],[Alice at 18 yrs.],[Kimberly],[Monica],[Decorator], [Park Avenue Couples],[Model],[Office Xmas Party Guest],[Dorothy's Maid],[Alice at 12 yrs.],[School Teacher],[Young Dorothy],[Park Avenue Couples],[Kate],[Muse], [Carol],[Circus Equestrian],[Nancy Brill].

Alice (1990) was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy,Romance. Annie Hall (1977) has the following cast :

[Tony's Date at Nightclub],[Pam],[Lacey Party Guest],[Janet],[Dorrie],[Allison],[Annie Hall],[Alvy's Classmate],[Grammy Hall],[Robin],[Alvy's Classmate],[Aunt Tessie],[Actress in Rehearsal],[Annie's Psychiatrist],[Street Stranger],[Street Stranger],[Alvy's Aunt], [School Teacher],[Alvy's Date Outside Theatre],[Alvy's Mom]. Annie Hall (1977) was produced by Gallo, Fred T., Greenhut, Robert(I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) is of genre type Drama.

Another Woman (1988) has the following cast :

[Hope],[Engagement Party Guest],[Waitress],[Cynthia],[Engagement Party Guest], [Clara],[Young Marion],[Young Claire],[Marion's Mother],[Laura],[Marion Post],[Young Clara],[Engagement Party Guest],[Little Marion],[Eleanor Banks].

Another Woman (1988) was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre type Comedy. Anything Else (2003) has the following cast :

[NA],[NA],[NA]. Anything Else (2003) was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] is in real life Cheng,Diane. [Alice] is in real life Farrow, Mia. [Joe's Daughter] is in real life Gallagher, Gina. [Alice at 18 yrs.] is in real life Graves, Kristy. [Kimberly] is in real life Henry, Alexi. [Monica] is in real life Hugot, Marceline. [Decorator] is in real life Kavner, Julie. [Park Avenue Couples] is in real life Long, Jodi. [Model] is in real life Macpherson, Elle. [Office Xmas Party Guest] is in real life Marie, Lisa (I). [Dorothy's Maid] is in real life Miley, Peggy. [Alice at12 yrs.] is in real life Miner, Rachel. [School Teacher] is in real life Moore,

Patience. [Young Dorothy] is in real life Nayber, Laurie. [Park Avenue Couples]is in real life O'Neill, Suzanne. [Kate] is in real life O'Sullivan Farrow, Dylan. [Muse] is in real life Peters, Bernadette. [Carol] is in real life Salinger,Diane. [Circus Equestrian] is in real life Schurmann, Katja. [Nancy Brill] is in real life Shepherd, Cybill.

[Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Lacey Party Guest] is in real life Frost, Jean Sarah. [Janet] is in real life Girard, Wendy. [Dorrie] is in real life Jones, Christine. [Allison] is in real life Kane, Carol. [Annie Hall] is in real life Keaton, Diane. [Alvy's Classmate] is in real life

Levitan, Amy. [Grammy Hall] is in real life Ludlam, Helen. [Robin] is in real life Margolin, Janet. [Alvy's Classmate] is in real life Mellinger, Susan. [Aunt Tessie] is in real life Novikoff, Rashel. [Actress in Rehearsal] is in real life Paris, Robin Mary. [Annie's Psychiatrist] is in real life Radburn, Veronica. [Street Stranger] is in real life Trueman, Paula. [Street Stranger] is in real life Tupper, Loretta. [Alvy's Aunt] is in real life Volner, Ruth. [School Teacher] is in real life Warwick, Margaretta. [Alvy's Date Outside Theatre] is in real life Weaver, Sigourney. [Alvy's Mom] is in real life Neuman, Joan. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [Waitress] is in real life Friedman, Dee Dee. [Cynthia] is in real life Grody, Kathryn. [Engagement Party Guest] is in real life Ivey, Dana. [Clara] is in real life Laslo, Mary. [Young Marion] is in real life Marx, Margaret. [Young Claire] is in real life McComb, Jennifer Lynn. [Marion's Mother] is in real life McGee, Caroline. [Laura] is in real life Plimpton, Martha. [Marion Post] is in real life Rowlands, Gena. [Young Clara] is in real life Schultz, Carol. [Engagement Party Guest] is in real life Spivak, Alice. [Little Marion] is in real life Sullivan, Heather. [Eleanor Banks] is in real life Whitaker, Jill. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana. [NA] is in real life Ricci, Christina.

Κείμενο γράφου σε split mode, χωρίς όρια στις μακροεντολές :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast: [Dr. Yang's Assistant],[Alice],[Joe's Daughter],[Alice at 18 yrs.],[Kimberly],[Monica],[Decorator], [Park Avenue Couples],[Model],[Office Xmas Party Guest],[Dorothy's Maid],[Alice at 12 yrs.],[School Teacher],[Young Dorothy],[Park Avenue Couples],[Kate],[Muse],[Carol],[Circus Equestrian],[Nancy Brill] and is of genre types Comedy,Drama,Fantasy,Romance and was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest],[Janet],[Dorrie],[Allison],[Annie Hall],[Alvy's Classmate],[Grammy Hall],[Robin],[Alvy's Classmate],[Aunt Tessie],[Actress in Rehearsal],[Annie's Psychiatrist],[Street Stranger],[Street Stranger],[Alvy's Aunt],[School Teacher],[Alvy's Date Outside Theatre],[Alvy's Mom] and is of genre types Comedy,Romance and was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress],[Cynthia],[Engagement Party Guest],[Clara],[Young Marion], [Young Claire], [Marion's Mother],[Laura],[Marion Post],[Young Clara],[Engagement Party Guest],[Little Marion],[Eleanor Banks] and is of genre types Drama and was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) has the following cast : [NA],[NA],[NA] and is of genre types Comedy and was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] is in real life Cheng,Diane. [Alice] is in real life Farrow, Mia. [Joe's Daughter] is in real life Gallagher, Gina. [Alice at 18 yrs.] is in real life Graves, Kristy. [Kimberly] is in real life Henry, Alexi. [Monica] is in real life Hugot, Marceline. [Decorator] is in real life Kavner, Julie. [Park Avenue Couples] is in real life Long, Jodi. [Model] is in real life Macpherson, Elle. [Office

Xmas Party Guest] is in real life Marie, Lisa (I). *[Dorothy's Maid]* is in real life Miley, Peggy. *[Alice at 12 yrs.]* is in real life Miner, Rachel. *[School Teacher]* is in real life Moore,

Patience. [Young Dorothy] is in real life Nayber, Laurie. *[Park Avenue Couples]* is in real life O'Neill, Suzanne. *[Kate]* is in real life O'Sullivan Farrow, Dylan. *[Muse]* is in real life Peters, Bernadette. *[Carol]* is in real life Salinger, Diane. *[Circus Equestrian]* is in real life Schurmann, Katja. *[Nancy Brill]* is in real life Shepherd, Cybill. *[Tony's Date at Nightclub]* is in real life Casey, Shaun. *[Pam]* is in real life Duvall, Shelley. *[Lacey Party Guest]* is in real life Frost, Jean Sarah. *[Janet]* is in real life Girard, Wendy. *[Dorrie]* is in real life Jones, Christine. *[Allison]* is in real life Kane, Carol. *[Annie Hall]* is in real life Keaton, Diane. *[Alvy's Classmate]* is in real life Levitan, Amy. *[Grammy Hall]* is in real life Ludlam, Helen. *[Robin]* is in real life Margolin, Janet. *[Alvy's Classmate]* is in real life Mellinger, Susan. *[Aunt Tessie]* is in real life Novikoff, Rashel. *[Actress in Rehearsal]* is in real life Paris, Robin Mary. *[Annie's Psychiatrist]* is in real life Radburn, Veronica. *[Street Stranger]* is in real life Trueman, Paula. *[Street Stranger]* is in real life Tupper, Loretta. *[Alvy's Aunt]* is in real life Volner, Ruth. *[School Teacher]* is in real life Warwick, Margaretta. *[Alvy's Date Outside Theatre]* is in real life Weaver, Sigourney. *[Alvy's Mom]* is in real life Neuman, Joan. *[Hope]* is in real life Farrow, Mia. *[Engagement Party Guest]* is in real life Frank, Janet. *[Waitress]* is in real life Friedman, Dee Dee. *[Cynthia]* is in real life Grody, Kathryn. *[Engagement Party Guest]* is in real life Ivey, Dana. *[Clara]* is in real life Laslo, Mary. *[Young Marion]* is in real life Marx, Margaret. *[Young Claire]* is in real life McComb, Jennifer Lynn. *[Marion's Mother]* is in real life McGee, Caroline. *[Laura]* is in real life Plimpton, Martha. *[Marion Post]* is in real life Rowlands, Gena. *[Young Clara]* is in real life Schultz, Carol. *[Engagement Party Guest]* is in real life Spivak, Alice. *[Little Marion]* is in real life Sullivan, Heather. *[Eleanor Banks]* is in real life Whitaker, Jill. *[NA]* is in real life Channing, Stockard. *[NA]* is in real life Krall, Diana. *[NA]* is in real life Ricci, Christina.

6.2.1.7 Μετάφραση με κενές τιμές

Σε περίπτωση που επιθυμούμε να επιχειρήσουμε να μεταφράσουμε κενές τιμές στην βάση, τότε υπάρχουν δυο περιπτώσεις αποτελεσμάτων. Και εξαρτώνται από τις μεταβλητές συστήματος. Για παράδειγμα, αν δοκιμάσουμε να βάλουμε στις μεταβλητές συστήματος του χρήστη στο πεδίο `NULL_WORD` την τιμή `<n/a>`, τότε για όλες τις παραπάνω παραμέτρους που αναφέρονται στην προηγούμενη παράγραφο, με την προσθήκη της ετικέτας σε ακμή προβολής

Ετικέτες ακμών προβολής (προηγούμενες +):

```
(hr, AWARD) = @DIR_NAME+' 's award is'+@BIRTHDATE
```

Επειδή το πεδίο `AWARD` της σχέσης `CAST` έχει κενές τιμές, τότε το αποτέλεσμα που λαμβάνουμε είναι :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast: [Dr. Yang's Assistant],[Alice] and is of genre types Comedy,Drama,Fantasy,Romance and was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam] and is of genre types Comedy,Romance and was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest] and is of genre types Drama and was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) has the following cast : [NA], [NA] and is of genre types Comedy and was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Engagement Party Guest] 's award is <n/a>. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana.

Βλέπουμε ότι για τις κενές τιμές, έχει αντικατασταθεί η κενή τιμή με την τιμή που θέσαμε εμείς στην μεταβλητή συστήματος. Για παράδειγμα *[Engagement Party Guest] 's award is <n/a>*.

Αν στις μεταβλητές συστήματος αφήσουμε στο πεδίο *NULL_WORD* (*in projections and joins*) κενή τιμή, αυτό σημαίνει για την μηχανή μετάφρασης πως πρέπει να αγνοήσει όλη την πρόταση και να μην την περιλάβει στην μετάφραση. Επομένως το αποτέλεσμα θα είναι ακριβώς το ίδιο με της προηγούμενης παραγράφου.

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast: [Dr. Yang's Assistant],[Alice] and is of genre types Comedy,Drama,Fantasy,Romance and was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam] and is of genre types Comedy,Romance and was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest] and is of genre types Drama and was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II).Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) has the following cast : [NA],[NA] and is of genre types Comedy and was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana.

Τα παραπάνω ισχύουν και για τις μακροεντολές. Αν για παράδειγμα, αντί για την ετικέτα

```
(hr, AWARD) = @DIR_NAME+' 's award is"+@BIRTHDATE
```

ορίζουμε μια μακροεντολή της μορφής :

```
DEFINE CAST_TUPLES_LIST as
[i < arityOf(@MOVIE_ID)]
{@MOVIECHAR[$i$]+' 's award is"+@AWARD[$i$]}
[i = arityOf(@MOVIE_ID)]
{@MOVIECHAR[$i$]+' 's award is "+@AWARD[$i$]+"."}

```

Τότε, ανάλογα με την μεταβλητή συστήματος *NULL_WORD* (*in macros*) = n/a στην δική μας περίπτωση, θα είχαμε :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast: [Dr. Yang's Assistant],[Alice] and is of genre types Comedy,Drama,Fantasy,Romance and was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam] and is of genre types Comedy,Romance and was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest] and is of genre types Drama and was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) has the following cast : [NA],[NA] and is of genre types Comedy and was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant]'s award is n/a.[Alice]'s award is n/a.[Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Tony's Date at Nightclub]'s award is n/a.[Pam]'s award is n/a.[Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Engagement Party Guest]'s award is n/a.[Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [NA]'s award is n/a.[NA]'s award is n/a.[NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana.

Αν, όπως και στην παραπάνω περίπτωση, θέταμε την μεταβλητή συστήματος *NULL_WORD* (*in macros*) κενή, τότε το αποτέλεσμα θα αγνοούσε και την μακροεντολή, με αποτέλεσμα :

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast: [Dr. Yang's Assistant],[Alice] and is of genre types Comedy,Drama,Fantasy,Romance and was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam] and is of genre types Comedy,Romance and was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest] and is of genre types Drama and was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II).Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) has the following cast : [NA],[NA] and is of genre types Comedy and was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana.

6.2.1.8 *Μετάφραση μετα- δεδομένων για ολόκληρη την βάση*

Στην περίπτωση που επιθυμούμε να δούμε όλα τα μετα- δεδομένα της βάσης και μόνο αυτά, τότε αν ακολουθήσουμε τις εξής ετικέτες και μακροεντολές,

Ετικέτες ακμών προβολής :

-

Ετικέτες ακμών σύνδεσης :

```
(DIRECTORS, DIRECTED_BY) = @DIR_NAME + "" + DIRECTED_BY_HA_LIST
```

```
(DIRECTED_BY, MOVIES) = @MOVIE_ID + "" + MOVIES_HA_LIST
```

```
(MOVIES, MGENRES) = @TITLE + "" + MGENRES_HA_LIST
```

```
(MOVIES, CAST) = @TITLE + "" + CAST_HA_LIST
(CAST, ACTORS) = @MOVIECHAR + "" + ACTORS_HA_LIST
(MOVIES, PRODUCED_BY) = @TITLE + "" + PRODUCED_BY_HA_LIST
(PRODUCED_BY, PRODUCERS) = @MOVIE_ID + "" + PRODUCERS_HA_LIST
```

Μακροεντολές :

```
DEFINE MOVIES_HA_LIST as
[i <= arityOf(@MOVIE_ID)]
{""}
```

```
DEFINE DIRECTED_BY_HA_LIST as
[i <= arityOf(@SUBDB)]
{""}
```

```
DEFINE CAST_HA_LIST as
[i <= arityOf(@MOVIE_ID)]
{@MOVIECHAR[$i$]+", "}
[i = arityOf(@MOVIE_ID)]
{""}
```

```
DEFINE ACTORS_HA_LIST as
[i <= arityOf(@AC_ID)]
{""}
```

```
DEFINE PRODUCED_BY_HA_LIST as
[i <= arityOf(@MOVIE_ID)]
{""}
```

```
DEFINE PRODUCERS_HA_LIST as
[i <= arityOf(@PROD_ID)]
{""}
```

```
DEFINE MGENRES_HA_LIST as
[i <= arityOf(@MOVIE_ID)]
```

```
{ "" }
```

```
DEFINE MGENRES_ATTR_LIST as
```

```
[i < arityOf(@MGENRES)]  
{@MGENRES[$i$]+", "  
[i = arityOf(@MGENRES)]  
{@MGENRES[$i$]+". "+"|nl|"} }
```

```
DEFINE PRODUCED_BY_ATTR_LIST as
```

```
[i < arityOf(@PRODUCED_BY)]  
{@PRODUCED_BY[$i$]+", "  
[i = arityOf(@PRODUCED_BY)]  
{@PRODUCED_BY[$i$]+". "+"|nl|"} }
```

```
DEFINE PRODUCERS_ATTR_LIST as
```

```
[i < arityOf(@PRODUCERS)]  
{@PRODUCERS[$i$]+", "  
[i = arityOf(@PRODUCERS)]  
{@PRODUCERS[$i$]+". "+"|nl|"} }
```

```
DEFINE DIRECTED_BY_ATTR_LIST as
```

```
[i < arityOf(@DIRECTED_BY)]  
{@DIRECTED_BY[$i$]+", "  
[i = arityOf(@DIRECTED_BY)]  
{@DIRECTED_BY[$i$]+". "+"|nl|"} }
```

```
DEFINE CAST_ATTR_LIST as
```

```
[i < arityOf(@CAST)]  
{@CAST[$i$]+", "  
[i = arityOf(@CAST)]  
{@CAST[$i$]+". "+"|nl|"} }
```

```
DEFINE DIRECTORS_ATTR_LIST as
```

```
[i < arityOf(@DIRECTORS)]  
{@DIRECTORS[$i$]+", "  
[i = arityOf(@DIRECTORS)]  
{@DIRECTORS[$i$]+". "+"|nl|"} }
```



```
[i = arityOf(@DIRECTORS)]
{@DIRECTORS[$i$]+". "+"|nl|"}

```

```
DEFINE MOVIES_ATTR_LIST as
[i < arityOf(@MOVIES)]
{@MOVIES[$i$]+", "}
[i = arityOf(@MOVIES)]
{@MOVIES[$i$]+". "+"|nl|"}

```

```
DEFINE ACTORS_ATTR_LIST as
[i < arityOf(@ACTORS)]
{@ACTORS[$i$]+", "}
[i = arityOf(@ACTORS)]
{@ACTORS[$i$]+". "+"|nl|"}

```

παίρνουμε ως αποτέλεσμα :

METADATA about DIRECTORS : SUBDB,BIRTHDATE,BIRTHLOCATION,DIR_ID,DIR_NAME.

METADATA about DIRECTED_BY : SUBDB,MOVIE_ID,DIR_ID.

METADATA about MOVIES : MOVIE_ID,TITLE,YEAR,DURATION,SUBDB.

METADATA about CAST : MOVIE_ID,AC_ID,AWARD,STARORDER,MOVIECHAR,SUBDB.

METADATA about PRODUCED_BY : MOVIE_ID,PROD_ID,PROD_POS,SUBDB.

METADATA about MGENRES : MOVIE_ID,GENRE,SUBDB.

METADATA about ACTORS : AC_ID,ACTOR_NAME,SEX,SUBDB.

METADATA about PRODUCERS : PROD_ID,PROD_NAME,SUBDB.

Σε κάθε ακμή σύνδεσης βάλαμε κενή μετάφραση, ούτως ώστε να μην εμφανιστεί τίποτα στον χρήστη για τις ακμές σύνδεσης, και για κάθε σχέση κατασκευάσαμε μια μακροεντολή που την περιγράφει. Στο τέλος της μακροεντολής προσθέσαμε και την έκφραση "|nl|", που όπως έχει αναφερθεί και παραπάνω, εκλαμβάνεται από την μηχανή μετάφρασης ως αλλαγή γραμμής, με αποτέλεσμα να έχουμε για κάθε σχέση μια πρόταση κάτω από την άλλη που περιγράφει τα μετα- δεδομένα της.

Θα μπορούσαμε, ακόμα, να μεταφράσουμε μετα- δεδομένα όποιων σχέσεων εμείς επιθυμούμε, απλά δηλώνοντας κενές τις μακροεντολές τους. Για παράδειγμα, με τις παραπάνω ετικέτες και μακροεντολές, αν αλλάζαμε τις εξής μακροεντολές,

```
DEFINE ACTORS_ATTR_LIST as [i <= arityOf (@ACTORS)] {""}  
DEFINE MOVIES_ATTR_LIST as [i <= arityOf (@MOVIES)] {""}
```

θα είχαμε ως αποτέλεσμα να μην μεταφραστούν οι συγκεκριμένες σχέσεις :

METADATA about DIRECTORS : SUBDB,BIRTHDATE,BIRTHLOCATION,DIR_ID,DIR_NAME.

METADATA about DIRECTED_BY : SUBDB,MOVIE_ID,DIR_ID.

METADATA about CAST : MOVIE_ID,AC_ID,AWARD,STARORDER,MOVIECHAR,SUBDB.

METADATA about PRODUCED_BY : MOVIE_ID,PROD_ID,PROD_POS,SUBDB.

METADATA about MGENRES : MOVIE_ID,GENRE,SUBDB.

METADATA about PRODUCERS : PROD_ID,PROD_NAME,SUBDB.

Σε παρόμοιο αποτέλεσμα, θα μπορούσαμε φυσικά να οδηγηθούμε και από την αλλαγή των βαρών στις ακμές.

6.3 Σύνοψη και επέκταση αποτελεσμάτων

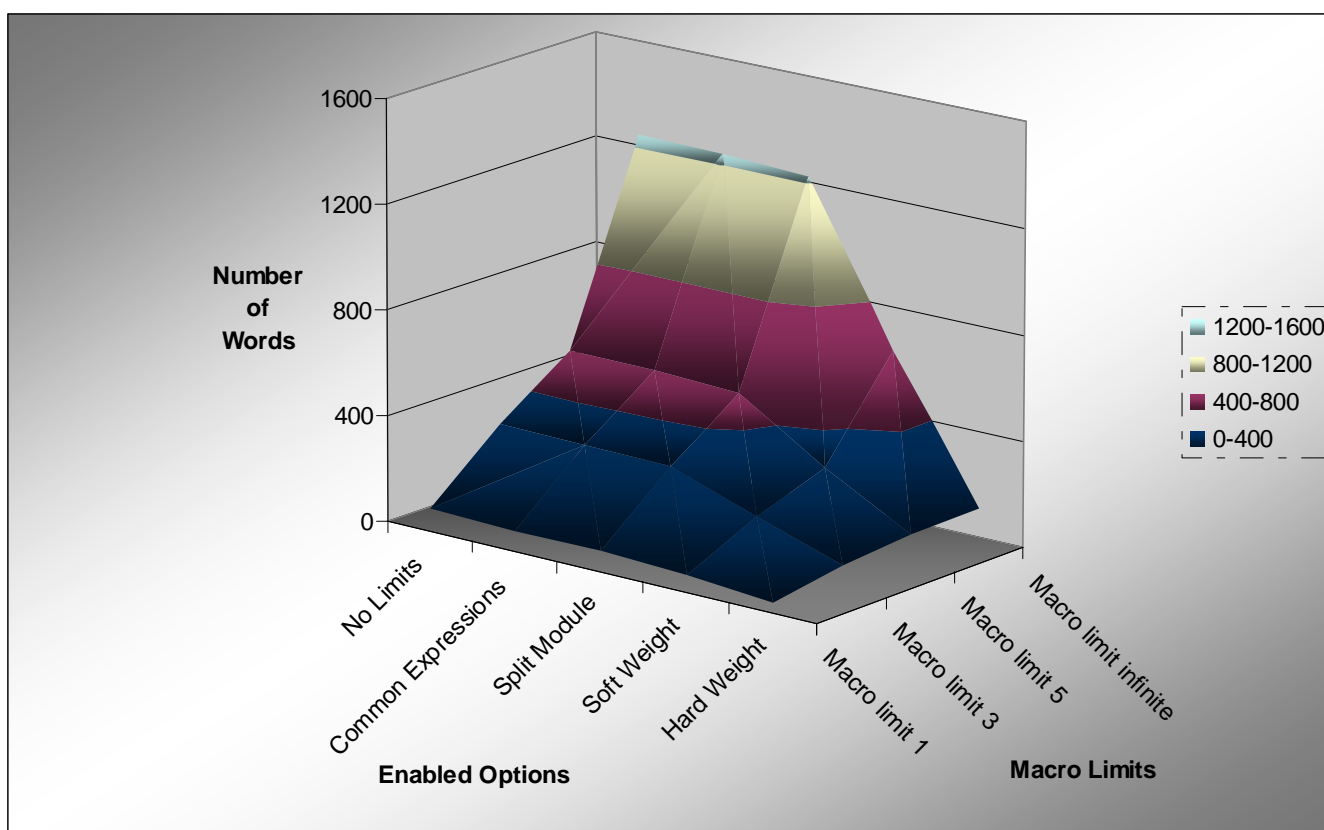
Είναι προφανές ότι δεν μπορεί να υπάρξει μια συγκεντρωτική πληροφόρηση όσον αφορά την ποιότητα από την άποψη της φυσικής γλώσσας του παραγόμενου κειμένου σε συνάρτηση με τις διάφορες επιλογές που δίδονται μέσω του γραφικού εργαλείου και υπαγορεύονται από την θεωρία των Précis Queries.

Οι επιλογές είναι τόσες πολλές, και οι διακυμάνσεις των αποτελεσμάτων καθώς αναμειγνύονται διαφορετικές επιλογές είναι τεράστιες, τόσες ώστε η ευμορφία του κειμένου να είναι ένα απολύτως υποκειμενικό θέμα το οποίο δεν επιδέχεται μονομερή και απόλυτα ορισμένη γνώμη.

Στο παρόν κεφάλαιο έγινε μια προσπάθεια αρίθμησης και παρουσίασης διαφόρων επιλογών για την δημιουργία κειμένου, όμως δεν είναι δυνατός ο ορισμός μιας, για παράδειγμα μεθοδολογίας, για την δημιουργία και παρουσίαση ικανοποιητικών, κατά το δοκούν του κάθε χρήστη, αποτελεσμάτων.

Εκτός από τις συμβουλές και τα συμπεράσματα που δίνονται παραπάνω για κάθε επιλογή που αναφέρεται συγκεκριμένα, εδώ παρουσιάζεται μια πειραματική δημιουργία κειμένου σε συνάρτηση με κάποιες παραμέτρους, που θεωρούνται κρίσιμες για την δημιουργία του αποτελέσματος και κυρίως έχουν να κάνουν με το μέγεθος του κειμένου και κάποιων κοινών εκφράσεων. Τα αποτελέσματα βρίσκονται στο ΠΑΡΑΡΤΗΜΑ Α', χάριν ευκολίας, μιας και περιέχουν και αρκετά αναλυτικές μεταφράσεις. Εδώ θα παρουσιαστεί μια γραφική παράσταση που περιέχει συνοπτικά τα αποτελέσματα.

Σημείωση : αυτό που παρουσιάζεται στην γραφική παράσταση και βρίσκεται αναλυτικά στο ΠΑΡΑΡΤΗΜΑ Α', δεν περιέχει αποτελέσματα όλων των δυνατών επιλογών, ούτε μπορεί να δηλώσει επακριβώς την συμπεριφορά του παραγόμενου κειμένου, μιας και αυτό έχει άμεση σχέση με τον γράφο, καθώς και τον τρόπο που θα χρησιμοποιηθούν οι επιλογές αυτές. Αυτό το παράδειγμα έχει σκοπό να δώσει μια γενική ιδέα όσον αφορά κάποιες επιλογές που θεωρούνται πιο κρίσιμες από άλλες, για ένα συγκεκριμένο κείμενο και γράφο.



Γραφική παράσταση πολυπλοκότητας/μήκους αποτελεσμάτων σε συνάρτηση με κάποιες από τις επιλογές της μηχανής μετάφρασης

Παρατηρούμε (με την βοήθεια και του παραρτήματος Α', όπου αναφέρονται αναλυτικά οι επιλογές και τα αποτελέσματα), ότι αν δεν εφαρμόσουμε κανένα όριο στο αποτέλεσμα, είτε μέσω βαρών είτε μέσω ορίων μακροεντολών, τότε χωρίς και καμία προσπάθεια δημιουργίας προτάσεων που θα μειώσουν τις “ξύλινες” εκφράσεις (μέσω εύρεσης κοινών εκφράσεων ή μέσω χρήσης του split module), καταλήγουμε με ένα μεγάλο κείμενο που πολλές φορές μεταφράζει το ίδιο πράγμα (για παράδειγμα το υποκείμενο – ΗΑ κάποιας σχέσης που μεταφράζονται οι ετικέτες προβολής χωρίς εύρεση κοινών εκφράσεων).

Δοκιμάζοντας να μορφοποιήσουμε τις ετικέτες κάποιων σχέσεων ούτως ώστε να χρησιμοποιηθεί η λειτουργικότητα της εύρεσης κοινών εκφράσεων, παρατηρούμε πως υπάρχει μια πολύ μικρή μείωση των λέξεων του κειμένου, όμως καταλήγουμε με πιο όμορφες και εκφράσεις που τείνουν στον φυσικό λόγο.

Αν ενεργοποιήσουμε και την λειτουργία του split module, αυτό γίνεται και στις ακμές σύνδεσης κάποιων σχέσεων, με αποτέλεσμα ένα πιο όμορφο και κατανοητό κείμενο.

Προχωρώντας ένα βήμα παραπέρα, αν αποκόψουμε κάποιες σχέσεις μέσω βαρών που ίσως να θεωρούμε τις πληροφορίες τους περιττές, τότε καταλήγουμε με ένα σαφώς πιο μικρό κείμενο που περιέχει λιγότερες πληροφορίες, όμως μπορεί να είναι επαρκείς για αυτό που πιθανώς χρειαζόμαστε, χωρίς να αναγκαστούμε να διαβάσουμε ένα μεγάλο κείμενο που πιθανώς να περιέχει περισσότερες, αλλά άχρηστες προς εμάς πληροφορίες.

Για κάθε μια από τις επιλογές παραπάνω, δηλώσαμε ένα όριο και για τις χρησιμοποιούμενες εντολές. Το όριο είναι κοινό για όλες τις μακροεντολές που χρησιμοποιούνται από την μηχανή μετάφρασης, αν και μπορούμε να χρησιμοποιήσουμε διαφορετικά όρια για διαφορετικές μακροεντολές. Οι επιλογές είναι πρακτικά άπειρες, μα ο σκοπός της συγκεκριμένης παραγράφου είναι μια θέαση κάποιου σεναρίου.

Παρατηρούμε ότι όσο μικραίνουμε το όριο των μεταφραζόμενων πλειάδων των μακροεντολών συρρικνώνεται δραστικά το κείμενο, το οποίο είναι αρκετά χρήσιμο αν δεν θέλουμε καταγισμό πληροφοριών για κάτι συγκεκριμένο, παρά κάποιες συνοπτικές πληροφορίες.

Μια διαφορετικού τύπου μείωση του μεγέθους του κειμένου θα μπορούσε να είναι η αποκοπή των προτάσεων που περιέχουν κενές (null) τιμές. Στις άνωθι εξεταζόμενες μεταφράσεις, οι null τιμές της βάσης αντικαθίστανται με μια τιμή “<n/a>”. Υπάρχει η δυνατότητα (αναφέρεται λεπτομερώς στον τρόπο χρήσης), να απορρίπτονται οι προτάσεις που περιέχουν

πεδία από την βάση τα οποία έχουν κενές τιμές. Με αυτόν τον τρόπο στην περίπτωση μας, θα είχαμε μεγάλη μείωση στο κείμενο, μιας και μια γρήγορα ματιά στα αποτελέσματα, θα δούμε πως υπάρχει το πεδίο AWARD το οποίο είναι κενό σε πολλές πλειάδες της βάσης στην σχέση CAST.

Περαιτέρω αλλαγές στο αποτέλεσμα της μετάφρασης θα μπορούσαν να έχουν η χρήση R° ή αλλού R^\bullet ανάλογα με το αποτέλεσμα που θα έδιναν. Αλλού, η χρήση μακροεντολών αντί για ετικέτες σε ακμές προβολής. Ή ακόμα και η χρήση μετα- δεδομένων της βάσης για την επισκόπησή της. Οι επιλογές είναι όσες και η φαντασία του χρήστη.

Γενικώς, η χρήση ορίων στις μακροεντολές, η τοποθέτηση βαρών στις ακμές, η αποκοπή πλειάδων με κενές τιμές, η χρήση του R° για την μετάφραση οδηγούν σε μικρότερου μεγέθους, συνοπτικά αποτελέσματα. Για πιο αναλυτικά αποτελέσματα μπορεί ο κάθε χρήστης να πειραματιστεί με την χρήση καταλληλότερων ορίων (ή και καθόλου), πάντα σε σχέση με την βάση που υπάρχει διαθέσιμη.

Χρήση μακροεντολών για τις ακμές προβολής επιτρέπει στον χρήστη μεγαλύτερη ευκολία στην δημιουργία κειμένου. Χρήση κατάλληλων κειμένων στις ακμές προβολής μπορεί να οδηγήσει σε οπτικά καλύτερα αποτελέσματα (μέσω της χρήσης του μηχανισμού εύρεσης κοινών εκφράσεων), πράγμα το οποίο ισχύει και για τις ακμές σύνδεσης, όταν αυτές συμμετέχουν σε ένα split mode.

Μετα- δεδομένα της βάσης επίσης μπορεί να φανούν χρήσιμα και εμφανίζονται μέσω εξειδικευμένων μακροεντολών (όπως έχουν αναφερθεί παραπάνω).

Για πιο εξειδικευμένες λειτουργίες του μηχανισμού μετάφρασης, υπάρχουν στο κεφάλαιο του τρόπου χρήσης της μηχανής και της γραφικής διεπαφής, καθώς και στα παραδείγματα μετάφρασης.

7

Επίλογος

Σε αυτό το κεφάλαιο θα συνοψίσουμε την παρουσίαση της διπλωματικής εργασίας.

7.1 Σύνοψη και συμπεράσματα

Ο σκοπός της παρούσας διπλωματικής ήταν η υλοποίηση ενός συστήματος παραγωγής ερωτήσεων *précis* (*précis queries*)[11]. Παραδείγματος χάριν, μια πιο ορθή νοηματικά απόκριση από την κλασική απάντηση που είναι σε μορφή πίνακα σε μια ερώτηση (*query*) που ρωτά για τον “Woody Allen” να είναι υπό μορφή ακόλουθων αποτελεσμάτων *précis* ερωτήσεων :

“Woody Allen was born on December 1, 1935 in Brooklyn, New York, USA.

As a director, Woody Allen’s work includes Match Point (2005), Melinda and Melinda (2004), Anything Else (2003). As an actor, Woody Allen’s work includes Hollywood Ending (2002), The Curse of the Jade Scorpion (2001).”

Για να παραχθεί ένα αποτέλεσμα αυτής της μορφής δημιουργήθηκε ένα γραφικό εργαλείο που ορίζει πιθανές παραμέτρους που μπορεί ο χρήστης να μεταβάλλει, ούτως ώστε να φέρει το αποτέλεσμα στα μέτρα του. Για την παραγωγή της απάντησης αναπτύχθηκε μια μηχανή μετάφρασης που μετατρέπει την βάση προς μετάφραση σε έναν γράφο και σύμφωνα με την διάσχιση του γράφου και τις παραμέτρους που έχει ορίσει ο χρήστης από το γραφικό εργαλείο, δημιουργεί και εμφανίζει την απάντηση στον χρήστη.

Από όλη την υλοποίηση διαφάνηκαν πολλά διαφορετικά συμπεράσματα.

Η χρήση μιας βάσης δεδομένων η οποία περιέχει σχέσεις με πολλά πεδία διευκολύνει ιδιαίτερα την μορφή της απάντησης, μιας και δίνει στον χρήστη μεγάλη ευελιξία όσον αφορά τον ορισμό προτύπων, ετικετών και μακροεντολών. Θεωρώντας ότι υπάρχουν διαφορετικά

προφίλ χρηστών, μια βάση με σχέσεις που περιέχουν πολλά πεδία μπορεί να εξυπηρετεί διαφορετικούς σκοπούς, όπως για παράδειγμα συνοπτικά αποτελέσματα σε κάποιο προφίλ περιλαμβάνοντας λίγα πεδία, ή εξαντλητικά αποτελέσματα περιλαμβάνοντας όλα τα πεδία των σχέσεων. Και ενδιάμεσα, ένας πεπερασμένος αριθμός διαφορετικών αποτελεσμάτων σε συνάρτηση με τον αριθμό των χρησιμοποιούμενων πεδίων. Στην βάση που χρησιμοποιήσαμε, ο αριθμός των πεδίων δεν ήταν ιδιαίτερα υψηλός, με αποτέλεσμα η προσφερόμενη ευελιξία όσον αφορά τα αποτελέσματα δεν ήταν μεγάλη. Παραδείγματος χάριν, η σχέση *DIRECTORS* περιείχε το πεδίο *DNNAME*, το οποίο περιείχε το όνομα και το επώνυμο του σκηνοθέτη στην μορφή *Allen, Woody*. Αυτό προφανώς και μετατρέπεται σε πρόβλημα ευελιξίας στην μετάφραση, μιας και η χρήση του υπαγορεύει να φαίνεται για πάντα έτσι, με αποτέλεσμα μια πιο “ξύλινη” γλώσσα και όχι κάποια περιγραφή κοντά σε φυσική γλώσσα, όπως και είναι ο σκοπός μας. Ο χωρισμός του πεδίου αυτού σε δυο άλλα, το ένα από τα οποία θα είχε το όνομα και το άλλο το επώνυμο θα προσφερόταν πιο πολύ, αν θέλαμε να προσθέσουμε ενδιάμεσά τους κάποια αλφαριθμητική έκφραση, ή ακόμα και να τα χρησιμοποιήσουμε σε διαφορετική σειρά, όπως *Woody Allen*, το οποίο μοιάζει περισσότερο με φυσική γλώσσα. Σε άλλη περίπτωση, θα μπορούσαμε να χρησιμοποιήσουμε το ένα εκ των δυο, χάριν συνοπτικών αποτελεσμάτων, χωρίς να χρειάζεται να φαίνονται και τα δυο στην μετάφραση. Αυτό ισχύει για τις ετικέτες αλλά και για τις μακροεντολές.

Στην βάση που είχαμε για την δημιουργία της μετάφρασης, υπήρχαν δυο συνδετικές σχέσεις, η *DIRECTED_BY* και η *PRODUCED_BY*, οι οποίες δεν είχαν μεταφραστικό ενδιαφέρον, μιας και περιείχαν μόνο τα κλειδιά που συνδέαν κάποιες σχέσεις μεταξύ τους. Είναι κοινή πρακτική να χρησιμοποιούνται τέτοιες συνδετικές σχέσεις στον σχεδιασμό βάσεων δεδομένων, διότι παρέχουν έναν τρόπο μείωσης του όγκου των πληροφοριών στην βάση δεδομένων. Όσον αφορά την μετάφραση όμως, λόγω της λειτουργίας του αλγορίθμου σύνδεσης σχέσεων, για κάθε πλειάδα μιας σχέσης- πατέρα R_i , τότε γίνεται μια σύνδεση (join) μεταξύ της σχέσης αυτής και της σχέσης- παιδιού R_j . Αυτό σημαίνει πως για μια συνδετική σχέση που παρεμβάλλεται ανάμεσα σε αυτές τις δυο, έστω R_x , γίνεται μια σύνδεση (join) μεταξύ $R_i \rightarrow R_x$, και μετά μια $R_x \rightarrow R_j$. Πιο περίπλοκα, αν από την σχέση R_i η σύνδεση με την R_j θα έδινε Y πλειάδες στην R_j , τότε αν παρεμβληθεί η R_x , τότε θα γινόταν μια σύνδεση $R_i \rightarrow R_x$ και θα έδινε Y πλειάδες. Για κάθε μια από αυτές λοιπόν, θα γινόταν μια σύνδεση $R_x \rightarrow R_j$. Παρατηρούμε έτσι ότι οι συνδέσεις (joins) που πραγματοποιούνται με αυτόν τον τρόπο να μην ακολουθούν πιστά τον αλγόριθμο και η διάσχιση του γράφου γίνεται κανονικά, όμως μιας και οι συνδετικές σχέσεις δεν υπολογίζονται στην μετάφραση, θα μπορούσαν να απορριφθούν και να μην υπολογιστούν οι επιπλέον συνδέσεις που μόνο αποτέλεσμα έχουν την αύξηση του χρόνου απόκρισης του μεταφραστή.

Επιπλέον, σε αυτό το σχήμα βάσης που χρησιμοποιήθηκε για την μετάφραση, (από www.imdb.com), δεν είναι δυνατή μια λειτουργία του μεταφραστή. Αυτή η λειτουργία αναφέρεται στην μετάφραση σε R° μορφή μια σχέση η οποία έχει εμφανιστεί ξανά στην μετάφραση και έχει μεταφραστεί σε R^\bullet . Αυτό θα συνέβαινε σε περίπτωση κατάληξης δυο διαφορετικών σχέσεων σε μια κοινή που θα έδιναν κάποιες κοινές πλειάδες. Η λειτουργία αυτή δεν διαφαίνεται στην παρούσα βάση, μιας και δεν υπάρχει στο σχήμα αυτό η προαναφερθείσα διάταξη. Σε παρόμοιες βάσεις που χρησιμοποιήθηκαν έγινε έλεγχος σε αυτήν την λειτουργία.

Γενικότερα, η λειτουργία της μηχανής μετάφρασης διαφαίνεται ότι ακολουθεί πιστά τις επιταγές της αντίστοιχης θεωρίας των *precis queries*. Τα αποτελέσματα είναι όπως αναμένονται σύμφωνα με την θεωρία, παρουσιάζοντας όλες τις διαφορετικές πτυχές των μεταφραστικών μεθόδων, χρησιμοποιώντας μακροεντολές, πρότυπα, μεταβλητές συστήματος. Η υλοποίηση αναλύει και προχωράει ένα βήμα περισσότερο την θεωρία, κατασκευάζοντας ένα βοηθητικό εργαλείο (η γραφική διεπαφή της εργασίας αυτής) προς χρήση από τους πιθανούς χρήστες του μεταφραστή, επιτρέποντας στον χρήστη την επισκόπηση της βάσης και την δημιουργία όλων των πιθανών παραμέτρων προς μετάφραση καθώς και εισάγοντας περισσότερους περιορισμούς προς χρήση, σύμφωνα με το κεφάλαιο του ελέγχου, όπου παρουσιάζονται όλες οι πιθανές επιλογές του χρήστη. Μια από αυτές είναι η χρήση περιορισμών των μακροεντολών που βρίσκονται πάνω στις ακμές σύνδεσης, όπου ανάλογα με τον αριθμητικό περιορισμό, μόνο αυτός ο αριθμός των πλειάδων της σχέσης στόχου θα μεταφραστούν, και αυτό χρησιμοποιείται αναδρομικά στις επόμενες σχέσεις προς μετάφραση, όπου χρησιμοποιούνται μόνο οι μεταφρασμένες πλειάδες. Αυτό εμφανίστηκε στην φάση της ανάπτυξης, όπου για μια βάση μπορεί το μέγεθος των αποτελεσμάτων να ήταν ιδιαίτερα αυξημένο, όπου οι επιλογές των βαρών δεν είχαν αντίκρισμα (μιας και ο νεο-εισαγόμενος περιορισμός αναφέρεται στις πλειάδες της σχέσης και όχι στην σχέση καθαυτή). Ο ορισμός και η χρήση πέντε διαφορετικών τύπων μακροεντολών επίσης επεκτείνουν και οριοθετούν την έννοια της μακροεντολής που ορίζεται στην θεωρία. Όλες οι επιλογές γίνονται από το γραφικό εργαλείο όπου οι χρήστες εισάγουν τις επιλογές τους και χρησιμοποιούνται από την μηχανή μετάφρασης για να δημιουργήσουν το αποτέλεσμα που θα επιστραφεί στον χρήστη. Επιπλέον, ο ορισμός λέξεων ή φράσεων που μπορεί να τοποθετηθούν στο αποτέλεσμα της μετάφρασης όταν το πεδίο είναι null (κενό), ή ακόμα και η απόρριψη ολόκληρης της πρότασης που περιέχει την κενή τιμή (στις ακμές προβολής), είναι επιπλέον χαρακτηριστικά που δημιουργήθηκαν κατά την ανάπτυξη της εργασίας.

7.2 Μελλοντικές επεκτάσεις

Ως συνέχιση της εργασίας αυτής, θα μπορούσε να θεωρηθεί η βελτιστοποίηση του αλγορίθμου σύνδεσης σχέσεων μεταξύ τους, λαμβάνοντας υπ' όψιν τις σχέσεις σύνδεσης που αναφέρθηκαν παραπάνω, οι οποίες περιέχουν τα κλειδιά των σχέσεων που συνδέουν και είναι κοινή πρακτική στον σχεδιασμό βάσεων, ούτως ώστε να μειωθούν οι συνδέσεις που πολλαπλασιάζονται με τον υπάρχοντα αλγόριθμο, επιτυγχάνοντας καλύτερες ταχύτητες μετάφρασης. Επιπλέον θα μπορούσε να δοκιμαστεί και μια μεταφορά του κώδικα της μηχανής μετάφρασης σε Java, για την σύγκρισή της με την παρούσα μορφή που έχει. Διαφορετικές βιβλιοθήκες γράφων θα μπορούσαν να χρησιμοποιηθούν για την εύρεση της πιο αποδοτικής μορφής για βελτιστοποίηση αποτελεσμάτων. Επίσης, όσον αφορά τα βάρη στις ακμές σύνδεσης των σχέσεων, αυτές συγκρίνονται με βάση την τρέχουσα τιμή του βάρους που έχουν. Σε διαφορετική περίπτωση, θα μπορούσε να υπολογίζεται το τρέχον βάρος της ακμής σύνδεσης ως το γινόμενο των προηγούμενων στον γράφο συνδέσεων που υπήρχαν στο μονοπάτι της μετάφρασης. Έτσι θα έχουμε ένα δυναμικό μοντέλο βαρών. Τέλος, θα μπορούσε να συζητηθεί και υλοποιηθεί μια τρίτη μορφή σύνδεσης σχέσεων στον γράφο της βάσης εκτός από τις δυο παρούσες (μοναδιαία ένωση και διαμεριστική ένωση), μια που αφορά μια σχέση στόχο R_j και πολλές σχέσεις πατέρες R_{i1}, \dots, R_{ik} . Με την ίδια λογική όπως στην διαμεριστική ένωση, θα μπορούσαν να ενωθούν οι ετικέτες των ακμών που στοχεύουν στην ίδια σχέση- παιδί για την αποφυγή επανάληψης των πλειάδων που είναι κοινές για όλες τις συνδέσεις των σχέσεων- πατέρων με την σχέση- παιδί.

8

Βιβλιογραφία

1. S. Agrawal, S. Chaudhuri, and G. Das. DBXplorer: A system for keyword-based search over relational databases. In ICDE, pp. 5-16, 2002.
2. I. Androutsopoulos, G.D. Ritchie, and P. Thanisch. Natural Language Interfaces to Databases - An Introduction. NL Eng., 1(1), pp. 29-81, 1995.
3. G. Bhalotia, A. Hulgeri, C. Nakhe, S. Chakrabarti, and S. Sudarshan. Keyword searching and browsing in databases using BANKS. In ICDE, pp. 431-440, 2002.
4. A. Dusterhoft, and B. Thalheim. Linguistic based search facilities in snowflake-like database schemes. DKE, 48, pp. 177-198, 2004.
5. D. Florescu, D. Kossmann, and I. Manolescu. Integrating keyword search into xml query processing. Computer Networks, 33(1-6), 2000.
6. L. Guo, F. Shao, C. Botev, and J. Shanmugasundaram. XRank: Ranked keyword search over XML documents. In SIGMOD, pp. 16-27, 2003.
7. L. R. Harris. User-Oriented Data Base Query with the ROBOT Natural Language Query System. VLDB 1977: 303-312.
8. V. Hristidis, L. Gravano, and Y. Papakonstantinou. Efficient IR-style keyword search over relational databases. In VLDB, pp. 850-861, 2003.
9. V. Hristidis, Y. Papakonstantinou, and A. Balmin. Keyword proximity search on XML graphs. In ICDE, pp. 367-378, 2003.
10. IBM. DB2 Text Information Extender. url: www.ibm.com/software/data/db2/extenders/textinformation/.
11. G. Koutrika, A. Simitsis, and Y. Ioannidis. Précis: The essence of a query answer. In ICDE, 2006.

12. U. Masermann, and G. Vossen. Design and implementation of a novel approach to keyword searching in relational databases. In ADBIS-DASFAA, pp. 171-184, 2000.
13. E. Metais, J. Meunier, and G. Levreau. Database Schema Design: A Perspective from Natural Language Techniques to Validation and View Integration. In ER, pp. 190-205, 2003.
14. E. Metais. Enhancing information systems management with natural language processing techniques. DKE, 41, pp. 247-272, 2002.
15. Microsoft. SQL Server 2000. url: <http://msdn.microsoft.com/library/>. 16. M. Minock. A Phrasal Approach to Natural Language Interfaces over Databases. In NLDB, pp. 181-191, 2005.
16. A. Motro. Baroque: A browser for relational databases. ACM Trans. Inf. Syst., 4(2), pp. 164-181, 1986.
17. A. Motro. Constructing queries from tokens. In SIGMOD, pp. 120-131, 1986.
18. Oracle. Oracle 9i Text. url: www.oracle.com/technology/products/text/.
19. A. Simitsis, and G. Koutrika. Pattern-Based Query Answering. In PaRMA, 2006.
20. E. Sneiders. Automated Question Answering Using Question Templates That Cover the Conceptual Model of the Database. In NLDB, pp. 235-239, 2002.
21. V.C. Storey, R.C. Goldstein, H. Ullrich. Naive Semantics to Support Automated Database Design. IEEE TKDE, 14(1), pp. 1-12, 2002.
22. V.C. Storey. Understanding and Representing Relationship Semantics in Database Design. In NLDB, pp. 79-90, 2001.
23. A. Toral, E. Noguera, F. Llopis, and R. Munoz. Improving Question Answering Using Named Entity Recognition. In NLDB, pp. 181-191, 2005.
24. P. Vassiliadis, A. Simitsis, P. Georgantas, M. Terrovitis, and S. Skiadopoulou. A Generic and Customizable Framework for the Design of ETL Scenarios. Information Systems, 30(7), pp. 492-525, 2005.
25. Q. Wang, C. Nass, and J. Hu. Natural Language Query vs. Keyword Search: Effects of Task Complexity on Search Performance, Participant Perceptions, and Preferences. In INTERACT, pp. 106-116, 2005.
26. A. Simitsis, G. Koutrika, Y. Ioannidis. Precis: From Unstructured Keywords as Queries to Structured Databases and Structured Text as Answers. Technical Report, 2007

ΠΑΡΑΡΤΗΜΑ Α

Σε αυτό το παράρτημα θα παρουσιάσουμε τα δεδομένα για τη σχήμα της παραγράφου 6.3 .

Ακολουθούν αναλυτικά τα κείμενα τα οποία βρίσκονται στην γραφική παράσταση, που αφορά την σχέση κάποιων επιλογών σε συνάρτηση με το μέγεθος σε λέξεις και την ποιότητα του παραγόμενου μεταφρασμένου κειμένου. Κάθε περαιτέρω επιλογή είναι προσθετική στις προηγούμενες (έτσι η χρήση του split mode αργότερα, περιέχει και τις κοινές εκφράσεις του προηγούμενου βήματος)

Μετάφραση χωρίς περιορισμούς

Χωρίς αλλαγές

- *Κανένα όριο στις μεταφραζόμενες πλειάδες των μακροεντολών.*
- *Καμία προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα (όπως για παράδειγμα δημιουργία ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων).*
- *Μη χρήση του split mode*
- *Κανένα βάρος στις ακμές*

ΛΕΞΕΙΣ : 1240

Allen, Woody 's birthlocation is Brooklyn, New York, USA. Allen, Woody was born on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title *Alice* (1990), it was directed on 1990 and its duration is 102 minutes. *Alice* (1990) is of genre types Comedy, Drama, Fantasy, Romance. *Alice* (1990) has the following cast : [Dr. Yang's Assistant], [Alice], [Joe's Daughter], [Alice at 18 yrs.], [Kimberly], [Monica], [Decorator], [Park Avenue Couples], [Model], [Office Xmas Party Guest], [Dorothy's Maid], [Alice at 12 yrs.], [School Teacher], [Young Dorothy], [Park Avenue Couples], [Kate], [Muse], [Carol], [Circus Equestrian], [Nancy Brill], [Hilda], [Helen], [Circus Aerialist], [Alice's Mother], [Penny], [Dorothy's Xmas Party Guest], [Nun]. *Alice* (1990) was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title *Annie Hall* (1977), it was directed on 1977 and its duration is 93 minutes. *Annie Hall* (1977) is of genre types Comedy, Romance. *Annie Hall* (1977) has the following cast : [Tony's Date at Nightclub], [Pam], [Lacey Party Guest], [Janet], [Dorrie], [Allison], [Annie Hall], [Alvy's Classmate], [Grammy Hall], [Robin], [Alvy's Classmate], [Aunt Tessie], [Actress in Rehearsal], [Annie's Psychiatrist], [Street Stranger], [Street Stranger], [Alvy's Aunt], [School Teacher], [Alvy's Date Outside Theatre], [Alvy's Mom], [Waitress at Health Food Restaurant]. *Annie Hall* (1977) was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title *Another Woman* (1988), it was directed on 1988 and its duration is 84 minutes. *Another Woman* (1988) is of genre type Drama. *Another Woman* (1988) has the following cast : [Hope], [Engagement Party Guest], [Waitress], [Cynthia], [Engagement Party Guest], [Clara], [Young Marion], [Young Claire], [Marion's Mother], [Laura], [Marion Post], [Young Clara], [Engagement Party Guest], [Little Marion], [Eleanor Banks]. *Another Woman* (1988) was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title *Anything Else* (2003), it was directed on 2003 and its duration is 96 minutes. *Anything Else* (2003) is of genre type Comedy. *Anything Else* (2003) has the following cast : [NA], [NA], [NA]. *Anything Else* (2003) was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Alice at 18 yrs.] 's award is <n/a>. [Kimberly] 's award is <n/a>. [Monica] 's award is <n/a>. [Decorator] 's award is <n/a>. [Park Avenue Couples] 's award is <n/a>. [Model] 's award is <n/a>. [Office Xmas Party Guest] 's award is <n/a>. [Dorothy's Maid] 's award is <n/a>. [Alice at 12 yrs.] 's award is <n/a>. [School Teacher] 's award is <n/a>. [Young Dorothy] 's award is <n/a>. [Park Avenue Couples] 's award is <n/a>. [Kate] 's award is <n/a>. [Muse] 's award is <n/a>. [Carol] 's award is <n/a>. [Circus Equestrian] 's award is <n/a>. [Nancy Brill] 's award is <n/a>. [Hilda] 's award is <n/a>. [Helen] 's award is <n/a>. [Circus Aerialist] 's award is <n/a>. [Alice's Mother] 's award is <n/a>. [Penny] 's

award is <n/a>. [Dorothy's Xmas Party Guest] 's award is <n/a>. [Nun] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Joe's Daughter] is in real life Gallagher, Gina. [Alice at 18 yrs.] is in real life Graves, Kristy. [Kimberly] is in real life Henry, Alexi. [Monica] is in real life Hugot, Marceline. [Decorator] is in real life Kavner, Julie. [Park Avenue Couples] is in real life Long, Jodi. [Model] is in real life Macpherson, Elle. [Office Xmas Party Guest] is in real life Marie, Lisa (I). [Dorothy's Maid] is in real life Miley, Peggy. [Alice at 12 yrs.] is in real life Miner, Rachel. [School Teacher] is in real life Moore, Patience. [Young Dorothy] is in real life Nayber, Laurie. [Park Avenue Couples] is in real life O'Neill, Suzanne. [Kate] is in real life O'Sullivan Farrow, Dylan. [Muse] is in real life Peters, Bernadette. [Carol] is in real life Salinger, Diane. [Circus Equestrian] is in real life Schurmann, Katja. [Nancy Brill] is in real life Shepherd, Cybill. [Hilda] is in real life Squibb, June. [Helen] is in real life Taylor, Holland. [Circus Aerialist] is in real life Thomas, Vanessa. [Alice's Mother] is in real life Verdon, Gwen. [Penny] is in real life Wallem, Linda. [Dorothy's Xmas Party Guest] is in real life Weston-Moran, Kim. [Nun] is in real life Stein, Mary (I). [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Janet] 's award is <n/a>. [Dorrie] 's award is <n/a>. [Allison] 's award is <n/a>. [Annie Hall] 's award is <n/a>. [Alvy's Classmate] 's award is <n/a>. [Grammy Hall] 's award is <n/a>. [Robin] 's award is <n/a>. [Alvy's Classmate] 's award is <n/a>. [Aunt Tessie] 's award is <n/a>. [Actress in Rehearsal] 's award is <n/a>. [Annie's Psychiatrist] 's award is <n/a>. [Street Stranger] 's award is <n/a>. [Street Stranger] 's award is <n/a>. [Alvy's Aunt] 's award is <n/a>. [School Teacher] 's award is <n/a>. [Alvy's Date Outside Theatre] 's award is <n/a>. [Alvy's Mom] 's award is <n/a>. [Waitress at Health Food Restaurant] 's award is <n/a>. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Lacey Party Guest] is in real life Frost, Jean Sarah. [Janet] is in real life Girard, Wendy. [Dorrie] is in real life Jones, Christine. [Allison] is in real life Kane, Carol. [Annie Hall] is in real life Keaton, Diane. [Alvy's Classmate] is in real life Levitan, Amy. [Grammy Hall] is in real life Ludlam, Helen. [Robin] is in real life Margolin, Janet. [Alvy's Classmate] is in real life Mellinger, Susan. [Aunt Tessie] is in real life Novikoff, Rashel. [Actress in Rehearsal] is in real life Paris, Robin Mary. [Annie's Psychiatrist] is in real life Radburn, Veronica. [Street Stranger] is in real life Trueman, Paula. [Street Stranger] is in real life Tupper, Loretta. [Alvy's Aunt] is in real life Volner, Ruth. [School Teacher] is in real life Warwick, Margaretta. [Alvy's Date Outside Theatre] is in real life Weaver, Sigourney. [Alvy's Mom] is in real life Neuman, Joan. [Waitress at Health Food Restaurant] is in real life Flippin, Lucy Lee. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>. [Cynthia] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Clara] 's

award is <n/a>. [Young Marion] 's award is <n/a>. [Young Claire] 's award is <n/a>. [Marion's Mother] 's award is <n/a>. [Laura] 's award is <n/a>. [Marion Post] 's award is <n/a>. [Young Clara] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Little Marion] 's award is <n/a>. [Eleanor Banks] 's award is <n/a>. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [Waitress] is in real life Friedman, Dee Dee. [Cynthia] is in real life Grody, Kathryn. [Engagement Party Guest] is in real life Ivey, Dana. [Clara] is in real life Laslo, Mary. [Young Marion] is in real life Marx, Margaret. [Young Claire] is in real life McComb, Jennifer Lynn. [Marion's Mother] is in real life McGee, Caroline. [Laura] is in real life Plimpton, Martha. [Marion Post] is in real life Rowlands, Gena. [Young Clara] is in real life Schultz, Carol. [Engagement Party Guest] is in real life Spivak, Alice. [Little Marion] is in real life Sullivan, Heather. [Eleanor Banks] is in real life Whitaker, Jill. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana. [NA] is in real life Ricci, Christina.

Κοινές εκφράσεις

- Κανένα όριο στις μεταφραζόμενες πλειάδες των μακροεντολών.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Μη χρήση του split mode
- Κανένα βάρος στις ακμές

ΛΕΞΕΙΣ : 1237

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter],[Alice at 18 yrs.],[Kimberly],[Monica],[Decorator],[Park Avenue Couples],[Model],[Office Xmas Party

Guest,*[Dorothy's Maid]*,*[Alice at 12 yrs.]*,*[School Teacher]*,*[Young Dorothy]*,*[Park Avenue Couples]*,*[Kate]*,*[Muse]*,*[Carol]*,*[Circus Equestrian]*,*[Nancy Brill]*,*[Hilda]*,*[Helen]*,*[Circus Aerialist]*,*[Alice's Mother]*,*[Penny]*,*[Dorothy's Xmas Party Guest]*,*[Nun]*. *Alice (1990)* was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title *Annie Hall (1977)*, it was directed on 1977 and its duration is 93 minutes. *Annie Hall (1977)* is of genre types *Comedy,Romance*. *Annie Hall (1977)* has the following cast : *[Tony's Date at Nightclub]*,*[Pam]*,*[Lacey Party Guest]*,*[Janet]*,*[Dorrie]*,*[Allison]*,*[Annie Hall]*,*[Alvy's Classmate]*,*[Grammy Hall]*,*[Robin]*,*[Alvy's Classmate]*,*[Aunt Tessie]*,*[Actress in Rehearsal]*,*[Annie's Psychiatrist]*,*[Street Stranger]*,*[Street Stranger]*,*[Alvy's Aunt]*,*[School Teacher]*,*[Alvy's Date Outside Theatre]*,*[Alvy's Mom]*,*[Waitress at Health Food Restaurant]*. *Annie Hall (1977)* was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title *Another Woman (1988)*, it was directed on 1988 and its duration is 84 minutes. *Another Woman (1988)* is of genre type *Drama*. *Another Woman (1988)* has the following cast : *[Hope]*,*[Engagement Party Guest]*,*[Waitress]*,*[Cynthia]*,*[Engagement Party Guest]*,*[Clara]*,*[Young Marion]*,*[Young Claire]*,*[Marion's Mother]*,*[Laura]*,*[Marion Post]*,*[Young Clara]*,*[Engagement Party Guest]*,*[Little Marion]*,*[Eleanor Banks]*. *Another Woman (1988)* was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title *Anything Else (2003)*, it was directed on 2003 and its duration is 96 minutes. *Anything Else (2003)* is of genre type *Comedy*. *Anything Else (2003)* has the following cast : *[NA]*,*[NA]*,*[NA]*. *Anything Else (2003)* was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). *[Dr. Yang's Assistant]* 's award is <n/a>. *[Alice]* 's award is <n/a>. *[Joe's Daughter]* 's award is <n/a>. *[Alice at 18 yrs.]* 's award is <n/a>. *[Kimberly]* 's award is <n/a>. *[Monica]* 's award is <n/a>. *[Decorator]* 's award is <n/a>. *[Park Avenue Couples]* 's award is <n/a>. *[Model]* 's award is <n/a>. *[Office Xmas Party Guest]* 's award is <n/a>. *[Dorothy's Maid]* 's award is <n/a>. *[Alice at 12 yrs.]* 's award is <n/a>. *[School Teacher]* 's award is <n/a>. *[Young Dorothy]* 's award is <n/a>. *[Park Avenue Couples]* 's award is <n/a>. *[Kate]* 's award is <n/a>. *[Muse]* 's award is <n/a>. *[Carol]* 's award is <n/a>. *[Circus Equestrian]* 's award is <n/a>. *[Nancy Brill]* 's award is <n/a>. *[Hilda]* 's award is <n/a>. *[Helen]* 's award is <n/a>. *[Circus Aerialist]* 's award is <n/a>. *[Alice's Mother]* 's award is <n/a>. *[Penny]* 's award is <n/a>. *[Dorothy's Xmas Party Guest]* 's award is <n/a>. *[Nun]* 's award is <n/a>. *[Dr. Yang's Assistant]* is in real life Cheng, Diane. *[Alice]* is in real life Farrow, Mia. *[Joe's Daughter]* is in real life Gallagher, Gina. *[Alice at 18 yrs.]* is in real life Graves, Kristy. *[Kimberly]* is in real life Henry, Alexi. *[Monica]* is in real life Hugot, Marceline. *[Decorator]* is in real life Kavner, Julie. *[Park Avenue Couples]* is in real life Long, Jodi.

[Model] is in real life Macpherson, Elle. *[Office Xmas Party Guest]* is in real life Marie, Lisa (I). *[Dorothy's Maid]* is in real life Miley, Peggy. *[Alice at 12 yrs.]* is in real life Miner, Rachel. *[School Teacher]* is in real life Moore, Patience. *[Young Dorothy]* is in real life Nayber, Laurie. *[Park Avenue Couples]* is in real life O'Neill, Suzanne. *[Kate]* is in real life O'Sullivan Farrow, Dylan. *[Muse]* is in real life Peters, Bernadette. *[Carol]* is in real life Salinger, Diane. *[Circus Equestrian]* is in real life Schurmann, Katja. *[Nancy Brill]* is in real life Shepherd, Cybill. *[Hilda]* is in real life Squibb, June. *[Helen]* is in real life Taylor, Holland. *[Circus Aerialist]* is in real life Thomas, Vanessa. *[Alice's Mother]* is in real life Verdon, Gwen. *[Penny]* is in real life Wallem, Linda. *[Dorothy's Xmas Party Guest]* is in real life Weston-Moran, Kim. *[Nun]* is in real life Stein, Mary (I). *[Tony's Date at Nightclub]* 's award is <n/a>. *[Pam]* 's award is <n/a>. *[Lacey Party Guest]* 's award is <n/a>. *[Janet]* 's award is <n/a>. *[Dorrie]* 's award is <n/a>. *[Allison]* 's award is <n/a>. *[Annie Hall]* 's award is <n/a>. *[Alvy's Classmate]* 's award is <n/a>. *[Grammy Hall]* 's award is <n/a>. *[Robin]* 's award is <n/a>. *[Alvy's Classmate]* 's award is <n/a>. *[Aunt Tessie]* 's award is <n/a>. *[Actress in Rehearsal]* 's award is <n/a>. *[Annie's Psychiatrist]* 's award is <n/a>. *[Street Stranger]* 's award is <n/a>. *[Street Stranger]* 's award is <n/a>. *[Alvy's Aunt]* 's award is <n/a>. *[School Teacher]* 's award is <n/a>. *[Alvy's Date Outside Theatre]* 's award is <n/a>. *[Alvy's Mom]* 's award is <n/a>. *[Waitress at Health Food Restaurant]* 's award is <n/a>. *[Tony's Date at Nightclub]* is in real life Casey, Shaun. *[Pam]* is in real life Duvall, Shelley. *[Lacey Party Guest]* is in real life Frost, Jean Sarah. *[Janet]* is in real life Girard, Wendy. *[Dorrie]* is in real life Jones, Christine. *[Allison]* is in real life Kane, Carol. *[Annie Hall]* is in real life Keaton, Diane. *[Alvy's Classmate]* is in real life Levitan, Amy. *[Grammy Hall]* is in real life Ludlam, Helen. *[Robin]* is in real life Margolin, Janet. *[Alvy's Classmate]* is in real life Mellinger, Susan. *[Aunt Tessie]* is in real life Novikoff, Rashel. *[Actress in Rehearsal]* is in real life Paris, Robin Mary. *[Annie's Psychiatrist]* is in real life Radburn, Veronica. *[Street Stranger]* is in real life Trueman, Paula. *[Street Stranger]* is in real life Tupper, Loretta. *[Alvy's Aunt]* is in real life Volner, Ruth. *[School Teacher]* is in real life Warwick, Margaretta. *[Alvy's Date Outside Theatre]* is in real life Weaver, Sigourney. *[Alvy's Mom]* is in real life Neuman, Joan. *[Waitress at Health Food Restaurant]* is in real life Flippin, Lucy Lee. *[Engagement Party Guest]* 's award is <n/a>. *[Waitress]* 's award is <n/a>. *[Cynthia]* 's award is <n/a>. *[Engagement Party Guest]* 's award is <n/a>. *[Clara]* 's award is <n/a>. *[Young Marion]* 's award is <n/a>. *[Young Claire]* 's award is <n/a>. *[Marion's Mother]* 's award is <n/a>. *[Laura]* 's award is <n/a>. *[Marion Post]* 's award is <n/a>. *[Young Clara]* 's award is <n/a>. *[Engagement Party Guest]* 's award is <n/a>. *[Little Marion]* 's award is <n/a>. *[Eleanor Banks]* 's award is <n/a>. *[Hope]* is in real life Farrow, Mia. *[Engagement Party Guest]* is in real life Frank, Janet. *[Waitress]*

is in real life Friedman, Dee Dee. [Cynthia] is in real life Grody, Kathryn. [Engagement Party Guest] is in real life Ivey, Dana. [Clara] is in real life Laslo, Mary. [Young Marion] is in real life Marx, Margaret. [Young Claire] is in real life McComb, Jennifer Lynn. [Marion's Mother] is in real life McGee, Caroline. [Laura] is in real life Plimpton, Martha. [Marion Post] is in real life Rowlands, Gena. [Young Clara] is in real life Schultz, Carol. [Engagement Party Guest] is in real life Spivak, Alice. [Little Marion] is in real life Sullivan, Heather. [Eleanor Banks] is in real life Whitaker, Jill. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana. [NA] is in real life Ricci, Christina.

Χρήση split mode

- *Κανένα όριο στις μεταφραζόμενες πλειάδες των μακροεντολών.*
- *Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.*
- *Χρήση του split mode*
- *Κανένα βάρος στις ακμές*

ΛΕΞΕΙΣ : 1223

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter],[Alice at 18 yrs.],[Kimberly],[Monica],[Decorator],[Park Avenue Couples],[Model],[Office Xmas Party Guest],[Dorothy's Maid],[Alice at 12 yrs.],[School Teacher],[Young Dorothy],[Park Avenue Couples],[Kate],[Muse],[Carol],[Circus Equestrian],[Nancy Brill],[Hilda],[Helen],[Circus Aerialist],[Alice's Mother],[Penny],[Dorothy's Xmas Party Guest],[Nun] and is of genre types Comedy,Drama,Fantasy,Romance and was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Annie Hall (1977), it was directed on 1977

and its duration is 93 minutes. *Annie Hall* (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest],[Janet],[Dorrie],[Allison],[Annie Hall],[Alvy's Classmate],[Grammy Hall],[Robin],[Alvy's Classmate],[Aunt Tessie],[Actress in Rehearsal],[Annie's Psychiatrist],[Street Stranger],[Street Stranger],[Alvy's Aunt],[School Teacher],[Alvy's Date Outside Theatre],[Alvy's Mom],[Waitress at Health Food Restaurant] and is of genre types Comedy,Romance and was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title *Another Woman* (1988), it was directed on 1988 and its duration is 84 minutes. *Another Woman* (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress],[Cynthia],[Engagement Party Guest],[Clara],[Young Marion],[Young Claire],[Marion's Mother],[Laura],[Marion Post],[Young Clara],[Engagement Party Guest],[Little Marion],[Eleanor Banks] and is of genre types Drama and was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title *Anything Else* (2003), it was directed on 2003 and its duration is 96 minutes. *Anything Else* (2003) has the following cast : [NA],[NA],[NA] and is of genre types Comedy and was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Alice at 18 yrs.] 's award is <n/a>. [Kimberly] 's award is <n/a>. [Monica] 's award is <n/a>. [Decorator] 's award is <n/a>. [Park Avenue Couples] 's award is <n/a>. [Model] 's award is <n/a>. [Office Xmas Party Guest] 's award is <n/a>. [Dorothy's Maid] 's award is <n/a>. [Alice at 12 yrs.] 's award is <n/a>. [School Teacher] 's award is <n/a>. [Young Dorothy] 's award is <n/a>. [Park Avenue Couples] 's award is <n/a>. [Kate] 's award is <n/a>. [Muse] 's award is <n/a>. [Carol] 's award is <n/a>. [Circus Equestrian] 's award is <n/a>. [Nancy Brill] 's award is <n/a>. [Hilda] 's award is <n/a>. [Helen] 's award is <n/a>. [Circus Aerialist] 's award is <n/a>. [Alice's Mother] 's award is <n/a>. [Penny] 's award is <n/a>. [Dorothy's Xmas Party Guest] 's award is <n/a>. [Nun] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Joe's Daughter] is in real life Gallagher, Gina. [Alice at 18 yrs.] is in real life Graves, Kristy. [Kimberly] is in real life Henry, Alexi. [Monica] is in real life Hugot, Marceline. [Decorator] is in real life Kavner, Julie. [Park Avenue Couples] is in real life Long, Jodi. [Model] is in real life Macpherson, Elle. [Office Xmas Party Guest] is in real life Marie, Lisa (I). [Dorothy's Maid] is in real life Miley, Peggy. [Alice at 12 yrs.] is in real life Miner, Rachel. [School Teacher] is in real life Moore, Patience. [Young Dorothy] is in real life Nayber, Laurie. [Park Avenue Couples] is in real life O'Neill, Suzanne. [Kate] is in real life O'Sullivan Farrow, Dylan. [Muse] is in real life Peters, Bernadette. [Carol] is in real life Salinger, Diane. [Circus Equestrian] is in real life Schurmann, Katja. [Nancy Brill] is in

real life Shepherd, Cybill. [Hilda] is in real life Squibb, June. [Helen] is in real life Taylor, Holland. [Circus Aerialist] is in real life Thomas, Vanessa. [Alice's Mother] is in real life Verdon, Gwen. [Penny] is in real life Wallem, Linda. [Dorothy's Xmas Party Guest] is in real life Weston-Moran, Kim. [Nun] is in real life Stein, Mary (I). [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Janet] 's award is <n/a>. [Dorrie] 's award is <n/a>. [Allison] 's award is <n/a>. [Annie Hall] 's award is <n/a>. [Alvy's Classmate] 's award is <n/a>. [Grammy Hall] 's award is <n/a>. [Robin] 's award is <n/a>. [Alvy's Classmate] 's award is <n/a>. [Aunt Tessie] 's award is <n/a>. [Actress in Rehearsal] 's award is <n/a>. [Annie's Psychiatrist] 's award is <n/a>. [Street Stranger] 's award is <n/a>. [Street Stranger] 's award is <n/a>. [Alvy's Aunt] 's award is <n/a>. [School Teacher] 's award is <n/a>. [Alvy's Date Outside Theatre] 's award is <n/a>. [Alvy's Mom] 's award is <n/a>. [Waitress at Health Food Restaurant] 's award is <n/a>. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Lacey Party Guest] is in real life Frost, Jean Sarah. [Janet] is in real life Girard, Wendy. [Dorrie] is in real life Jones, Christine. [Allison] is in real life Kane, Carol. [Annie Hall] is in real life Keaton, Diane. [Alvy's Classmate] is in real life Levitan, Amy. [Grammy Hall] is in real life Ludlam, Helen. [Robin] is in real life Margolin, Janet. [Alvy's Classmate] is in real life Mellinger, Susan. [Aunt Tessie] is in real life Novikoff, Rashel. [Actress in Rehearsal] is in real life Paris, Robin Mary. [Annie's Psychiatrist] is in real life Radburn, Veronica. [Street Stranger] is in real life Trueman, Paula. [Street Stranger] is in real life Tupper, Loretta. [Alvy's Aunt] is in real life Volner, Ruth. [School Teacher] is in real life Warwick, Margaretta. [Alvy's Date Outside Theatre] is in real life Weaver, Sigourney. [Alvy's Mom] is in real life Neuman, Joan. [Waitress at Health Food Restaurant] is in real life Flippin, Lucy Lee. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>. [Cynthia] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Clara] 's award is <n/a>. [Young Marion] 's award is <n/a>. [Young Claire] 's award is <n/a>. [Marion's Mother] 's award is <n/a>. [Laura] 's award is <n/a>. [Marion Post] 's award is <n/a>. [Young Clara] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Little Marion] 's award is <n/a>. [Eleanor Banks] 's award is <n/a>. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [Waitress] is in real life Friedman, Dee Dee. [Cynthia] is in real life Grody, Kathryn. [Engagement Party Guest] is in real life Ivey, Dana. [Clara] is in real life Laslo, Mary. [Young Marion] is in real life Marx, Margaret. [Young Claire] is in real life McComb, Jennifer Lynn. [Marion's Mother] is in real life McGee, Caroline. [Laura] is in real life Plimpton, Martha. [Marion Post] is in real life Rowlands, Gena. [Young Clara] is in real life Schultz, Carol. [Engagement Party Guest] is in real life Spivak, Alice. [Little Marion] is in real life Sullivan,

Heather. [Eleanor Banks] is in real life Whitaker, Jill. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana. [NA] is in real life Ricci, Christina.

Χρήση βαρών

- Κανένα όριο στις μεταφραζόμενες πλειάδες των μακροεντολών.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του split mode
- Αποκοπή μέσω βαρών κάποιων ακραίων σχέσεων (βάρη μόνο σε ακμές σύνδεσης), τις ACTORS, PRODUCED_BY

ΛΕΞΕΙΣ : 636

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter],[Alice at 18 yrs.],[Kimberly],[Monica],[Decorator],[Park Avenue Couples],[Model],[Office Xmas Party Guest],[Dorothy's Maid],[Alice at 12 yrs.],[School Teacher],[Young Dorothy],[Park Avenue Couples],[Kate],[Muse],[Carol],[Circus Equestrian],[Nancy Brill],[Hilda],[Helen],[Circus Aerialist],[Alice's Mother],[Penny],[Dorothy's Xmas Party Guest],[Nun] and is of genre types Comedy,Drama,Fantasy,Romance and was produced by Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest],[Janet],[Dorrie],[Allison],[Annie Hall],[Alvy's Classmate],[Grammy Hall],[Robin],[Alvy's Classmate],[Aunt Tessie],[Actress in Rehearsal],[Annie's Psychiatrist],[Street Stranger],[Street Stranger],[Alvy's Aunt],[School Teacher],[Alvy's Date Outside Theatre],[Alvy's Mom],[Waitress at Health Food Restaurant] and is of genre types Comedy,Romance and was produced by Regarding the movie with title Another Woman

(1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress],[Cynthia],[Engagement Party Guest],[Clara],[Young Marion],[Young Claire],[Marion's Mother],[Laura],[Marion Post],[Young Clara],[Engagement Party Guest],[Little Marion],[Eleanor Banks] and is of genre types Drama and was produced by Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) has the following cast : [NA],[NA],[NA] and is of genre types Comedy and was produced by [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Alice at 18 yrs.] 's award is <n/a>. [Kimberly] 's award is <n/a>. [Monica] 's award is <n/a>. [Decorator] 's award is <n/a>. [Park Avenue Couples] 's award is <n/a>. [Model] 's award is <n/a>. [Office Xmas Party Guest] 's award is <n/a>. [Dorothy's Maid] 's award is <n/a>. [Alice at 12 yrs.] 's award is <n/a>. [School Teacher] 's award is <n/a>. [Young Dorothy] 's award is <n/a>. [Park Avenue Couples] 's award is <n/a>. [Kate] 's award is <n/a>. [Muse] 's award is <n/a>. [Carol] 's award is <n/a>. [Circus Equestrian] 's award is <n/a>. [Nancy Brill] 's award is <n/a>. [Hilda] 's award is <n/a>. [Helen] 's award is <n/a>. [Circus Aerialist] 's award is <n/a>. [Alice's Mother] 's award is <n/a>. [Penny] 's award is <n/a>. [Dorothy's Xmas Party Guest] 's award is <n/a>. [Nun] 's award is <n/a>. [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Janet] 's award is <n/a>. [Dorrie] 's award is <n/a>. [Allison] 's award is <n/a>. [Annie Hall] 's award is <n/a>. [Alvy's Classmate] 's award is <n/a>. [Grammy Hall] 's award is <n/a>. [Robin] 's award is <n/a>. [Alvy's Classmate] 's award is <n/a>. [Aunt Tessie] 's award is <n/a>. [Actress in Rehearsal] 's award is <n/a>. [Annie's Psychiatrist] 's award is <n/a>. [Street Stranger] 's award is <n/a>. [Street Stranger] 's award is <n/a>. [Alvy's Aunt] 's award is <n/a>. [School Teacher] 's award is <n/a>. [Alvy's Date Outside Theatre] 's award is <n/a>. [Alvy's Mom] 's award is <n/a>. [Waitress at Health Food Restaurant] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>. [Cynthia] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Clara] 's award is <n/a>. [Young Marion] 's award is <n/a>. [Young Claire] 's award is <n/a>. [Marion's Mother] 's award is <n/a>. [Laura] 's award is <n/a>. [Marion Post] 's award is <n/a>. [Young Clara] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Little Marion] 's award is <n/a>. [Eleanor Banks] 's award is <n/a>. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] 's award is <n/a>.

- Κανένα όριο στις μεταφραζόμενες πλειάδες των μακροεντολών.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του *split mode*
- Αποκοπή μέσω βαρών όλων των σχέσεων από το 1^ο βήμα (κρατείται η μετάφραση μέχρι την σχέση MOVIES)

ΛΕΞΕΙΣ : 107

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes.

Όριο πέντε μεταφραζόμενων πλειάδων στις μακροεντολές

Χωρίς αλλαγές

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 5.
- Καμία προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα (όπως για παράδειγμα δημιουργία ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων).
- Μη χρήση του *split mode*
- Κανένα βάρος στις ακμές

ΛΕΞΕΙΣ : 506

Allen, Woody 's birthlocation is Brooklyn, New York, USA. Allen, Woody was born on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter],[Alice at 18 yrs.],[Kimberly]. Alice (1990) was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A.. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy,Romance. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest],[Janet],[Dorrie]. Annie Hall (1977) was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) is of genre type Drama. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress],[Cynthia],[Engagement Party Guest]. Another Woman (1988) was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre type Comedy. Anything Else (2003) has the following cast : [NA],[NA],[NA]. Anything Else (2003) was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Alice at 18 yrs.] 's award is <n/a>. [Kimberly] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Joe's Daughter] is in real life Gallagher, Gina. [Alice at 18 yrs.] is in real life Graves, Kristy. [Kimberly] is in real life Henry, Alexi. [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Janet] 's award is <n/a>. [Dorrie] 's award is <n/a>. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Lacey Party Guest] is in real life Frost, Jean Sarah. [Janet] is in real life Girard, Wendy. [Dorrie] is in real life Jones, Christine. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>. [Cynthia] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [Waitress] is in real life Friedman, Dee Dee. [Cynthia] is in real

life Grody, Kathryn. [Engagement Party Guest] is in real life Ivey, Dana. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana. [NA] is in real life Ricci, Christina.

Κοινές εκφράσεις

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 5.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Μη χρήση του split mode
- Κανένα βάρος στις ακμές

ΑΕΞΕΙΣ : 503

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy,Romance. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter],[Alice at 18 yrs.],[Kimberly]. Alice (1990) was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A.. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy,Romance. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest],[Janet],[Dorrie]. Annie Hall (1977) was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) is of genre type Drama. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress],[Cynthia],[Engagement Party Guest]. Another Woman (1988) was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) is of genre type

Comedy. *Anything Else* (2003) has the following cast : [NA],[NA],[NA]. *Anything Else* (2003) was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Alice at 18 yrs.] 's award is <n/a>. [Kimberly] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Joe's Daughter] is in real life Gallagher, Gina. [Alice at 18 yrs.] is in real life Graves, Kristy. [Kimberly] is in real life Henry, Alexi. [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Janet] 's award is <n/a>. [Dorrie] 's award is <n/a>. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Lacey Party Guest] is in real life Frost, Jean Sarah. [Janet] is in real life Girard, Wendy. [Dorrie] is in real life Jones, Christine. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>. [Cynthia] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [Waitress] is in real life Friedman, Dee Dee. [Cynthia] is in real life Grody, Kathryn. [Engagement Party Guest] is in real life Ivey, Dana. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana. [NA] is in real life Ricci, Christina.

Χρήση split mode

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 5.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του split mode
- Κανένα βάρος στις ακμές

ΛΕΞΕΙΣ : 489

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast : [Dr. Yang's

Assistant],[Alice],[Joe's Daughter],[Alice at 18 yrs.],[Kimberly] and is of genre types Comedy,Drama,Fantasy,Romance and was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H., Martin, Jane Read, Reilly, Thomas A.. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest],[Janet],[Dorrie] and is of genre types Comedy,Romance and was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H., Rollins, Jack (II). Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress],[Cynthia],[Engagement Party Guest] and is of genre types Drama and was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A., Robin, Helen, Rollins, Jack (II). Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) has the following cast : [NA],[NA],[NA] and is of genre types Comedy and was produced by Aronson, Letty, Joffe, Charles H., Rollins, Jack (II). [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Alice at 18 yrs.] 's award is <n/a>. [Kimberly] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Joe's Daughter] is in real life Gallagher, Gina. [Alice at 18 yrs.] is in real life Graves, Kristy. [Kimberly] is in real life Henry, Alexi. [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Janet] 's award is <n/a>. [Dorrie] 's award is <n/a>. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Lacey Party Guest] is in real life Frost, Jean Sarah. [Janet] is in real life Girard, Wendy. [Dorrie] is in real life Jones, Christine. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>. [Cynthia] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [Waitress] is in real life Friedman, Dee Dee. [Cynthia] is in real life Grody, Kathryn. [Engagement Party Guest] is in real life Ivey, Dana. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] is in real life Channing, Stockard. [NA] is in real life Krall, Diana. [NA] is in real life Ricci, Christina.

Χρήση βαρών

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 5.

- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του *split mode*
- Αποκοπή μέσω βαρών κάποιων ακραίων σχέσεων (βάρη μόνο σε ακμές σύνδεσης), τις *ACTORS, PRODUCED_BY*

ΛΕΞΕΙΣ : 282

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter],[Alice at 18 yrs.],[Kimberly] and is of genre types Comedy,Drama,Fantasy,Romance. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest],[Janet],[Dorrie] and is of genre types Comedy,Romance. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress],[Cynthia],[Engagement Party Guest] and is of genre types Drama. Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes. Anything Else (2003) has the following cast : [NA],[NA],[NA] and is of genre types Comedy. [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Alice at 18 yrs.] 's award is <n/a>. [Kimberly] 's award is <n/a>. [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Janet] 's award is <n/a>. [Dorrie] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>. [Cynthia] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [NA] 's award is <n/a>. [NA] 's award is <n/a>. [NA] 's award is <n/a>.

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 5.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του *split mode*
- Αποκοπή μέσω βαρών όλων των σχέσεων από το 1^ο βήμα (κρατείται η μετάφραση μέχρι την σχέση MOVIES)

ΛΕΞΕΙΣ : 107

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988),

Anything Else (2003). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Regarding the movie with title Anything Else (2003), it was directed on 2003 and its duration is 96 minutes.

Όριο τριών μεταφραζόμενων πλειάδων στις μακροεντολές

Χωρίς αλλαγές

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 3.
- Καμία προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα (όπως για παράδειγμα δημιουργία ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων).
- Μη χρήση του *split mode*
- Κανένα βάρος στις ακμές

ΛΕΞΕΙΣ : 316

Allen, Woody 's birthlocation is Brooklyn, New York, USA. Allen, Woody was born on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter]. Alice (1990) was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H.. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy,Romance. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest]. Annie Hall (1977) was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H.. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) is of genre type Drama. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress]. Another Woman (1988) was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A.. [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Joe's Daughter] is in real life Gallagher, Gina. [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Lacey Party Guest] is in real life Frost, Jean Sarah. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [Waitress] is in real life Friedman, Dee Dee.

Κοινές εκφράσεις

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 3.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Μη χρήση του *split mode*
- Κανένα βάρος στις ακμές

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre types Comedy,Drama,Fantasy. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter]. Alice (1990) was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H.. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) is of genre types Comedy,Romance. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest]. Annie Hall (1977) was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H.. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) is of genre type Drama. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress]. Another Woman (1988) was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A.. [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Joe's Daughter] is in real life Gallagher, Gina. [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Lacey Party Guest] is in real life Frost, Jean Sarah. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [Waitress] is in real life Friedman, Dee Dee.

Χρήση split mode

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 3.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του split mode
- Κανένα βάρος στις ακμές

ΛΕΞΕΙΣ : 303

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter] and is of genre types Comedy,Drama,Fantasy and was produced by Greenhut, Robert (I), Hartwick, Joseph, Joffe, Charles H.. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest] and is of genre types Comedy,Romance and was produced by Gallo, Fred T., Greenhut, Robert (I), Joffe, Charles H.. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress] and is of genre types Drama and was produced by Greenhut, Robert (I), Joffe, Charles H., Reilly, Thomas A. [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane. [Alice] is in real life Farrow, Mia. [Joe's Daughter] is in real life Gallagher, Gina. [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Tony's Date at Nightclub] is in real life Casey, Shaun. [Pam] is in real life Duvall, Shelley. [Lacey Party Guest] is in real life Frost, Jean Sarah. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>. [Hope] is in real life Farrow, Mia. [Engagement Party Guest] is in real life Frank, Janet. [Waitress] is in real life Friedman, Dee Dee.

Χρήση βαρών

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 3.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του *split mode*
- Αποκοπή μέσω βαρών κάποιων ακραίων σχέσεων (βάρη μόνο σε ακμές σύνδεσης), τις ACTORS, PRODUCED_BY

ΛΕΞΕΙΣ : 190

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast : [Dr. Yang's Assistant],[Alice],[Joe's Daughter] and is of genre types Comedy,Drama,Fantasy. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Annie Hall (1977) has the following cast : [Tony's Date at Nightclub],[Pam],[Lacey Party Guest] and is of genre types Comedy,Romance. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes. Another Woman (1988) has the following cast : [Hope],[Engagement Party Guest],[Waitress] and is of genre types Drama. [Dr. Yang's Assistant] 's award is <n/a>. [Alice] 's award is <n/a>. [Joe's Daughter] 's award is <n/a>. [Tony's Date at Nightclub] 's award is <n/a>. [Pam] 's award is <n/a>. [Lacey Party Guest] 's award is <n/a>. [Engagement Party Guest] 's award is <n/a>. [Waitress] 's award is <n/a>.

Χρήση επιπλέον βαρών

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 3.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του *split mode*
- Αποκοπή μέσω βαρών όλων των σχέσεων από το 1^ο βήμα (κρατείται η μετάφραση μέχρι την σχέση MOVIES)

ΛΕΞΕΙΣ : 86

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990),

Annie Hall (1977),

Another Woman (1988). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Regarding the movie with title Annie Hall (1977), it was directed on 1977 and its duration is 93 minutes. Regarding the movie with title Another Woman (1988), it was directed on 1988 and its duration is 84 minutes.

Όριο μιας μεταφραζόμενης πλειάδας στις μακροεντολές

Χωρίς αλλαγές

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 1.
- Καμία προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα (όπως για παράδειγμα δημιουργία ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων).
- Μη χρήση του split mode
- Κανένα βάρος στις ακμές

ΛΕΞΕΙΣ : 86

Allen, Woody 's birthlocation is Brooklyn, New York, USA. Allen, Woody was born on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre type Comedy. Alice (1990) has the following cast : [Dr. Yang's Assistant]. Alice (1990) was produced by Greenhut, Robert (I). [Dr. Yang's Assistant] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane.

Κοινές εκφράσεις

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 1.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Μη χρήση του split mode
- Κανένα βάρος στις ακμές

ΛΕΞΕΙΣ : 82

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) is of genre type Comedy. Alice (1990) has the following cast : [Dr. Yang's Assistant]. Alice (1990) was produced by Greenhut, Robert (I). [Dr. Yang's Assistant] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane.

Χρήση split mode

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 1.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του split mode
- Κανένα βάρος στις ακμές

ΛΕΞΕΙΣ : 80

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast : [Dr. Yang's Assistant] and is of genre types Comedy and was produced by Greenhut, Robert (I). [Dr. Yang's Assistant] 's award is <n/a>. [Dr. Yang's Assistant] is in real life Cheng, Diane.

Χρήση βαρών

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 1.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του split mode
- Αποκοπή μέσω βαρών κάποιων ακραίων σχέσεων (βάρη μόνο σε ακμές σύνδεσης), τις ACTORS, PRODUCED_BY

ΛΕΞΕΙΣ : 64

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes. Alice (1990) has the following cast : [Dr. Yang's Assistant] and is of genre types Comedy. [Dr. Yang's Assistant] 's award is <n/a>.

Χρήση επιπλέον βαρών

- Όριο στις μεταφραζόμενες πλειάδες των μακροεντολών ίσο με 1.
- Προσπάθεια διαμόρφωσης του κειμένου για πιο εύμορφα αποτελέσματα με την χρήση ετικετών που θα βρεθούν κοινές εκφράσεις για την αποφυγή επαναλαμβανόμενων εκφράσεων.
- Χρήση του split mode
- Αποκοπή μέσω βαρών όλων των σχέσεων από το 1^ο βήμα (κρατείται η μετάφραση μέχρι την σχέση MOVIES)

ΛΕΞΕΙΣ : 41

Allen, Woody was born in Brooklyn, New York, USA, on December 1, 1935. Allen, Woody has directed the following movies :

Alice (1990). Regarding the movie with title Alice (1990), it was directed on 1990 and its duration is 102 minutes.