Σχεσιακός Λογισμός (Relational Calculus)

- γλώσσα βασισμένη στον Κατηγορηματικό Λογισμό 1ης Τάξης (1st-order predicate calculus)
- οι περισσότερες γλώσσες επερώτησης σχεσιακών βάσεων δεδομένων βασίζονται στον Σχεσιακό Λογισμό
- έχει δύο μορφές:
  - σχεσιακός λογισμός πεδίων (domain relational calculus)
  - σχεσιακός λογισμός πλειάδων (tuple relational calculus)
- ο σχεσιακός λογισμός και η σχεσιακή άλγεβρα έχουν την ίδια εκφραστική δύναμη
- μια γλώσσα επερωτήσεων λέγεται **πλήρης** αν έχει την ίδια εκφραστική δύναμη με την Σ.Α. ή το Σ.Λ.

Προτάσεις του Σ.Λ.

- Οι προτάσεις του Σ.Λ. δηλώνουν σχέσεις (πιθανά μη-πεπερασμένες).
- Κάθε πρόταση χρησιμοποιεί ένα σύνολο μεταβλητών οι οποίες διακρίνονται σε ελεύθερες (free) και δεσμευμένες (bound). Η ίδια μεταβλητή μπορεί να εμφανίζεται ως ελεύθερη και ως δεσμευμένη στην ίδια πρόταση.
- Το σχεσιακό σχήμα για μια πρόταση του Σ.Λ. είναι ένα σύνολο γνωρισμάτων τα οποία αντιστοιχούν στις ελεύθερες μεταβλητές της πρότασης.

Οι καλά-δομημένες προτάσεις του Σ.Λ. ακολουθούν τον παρακάτω αναδρομικό ορισμό:

- οι ατομικές προτάσεις είναι καλά-δομημένες προτάσεις:
  1. για κάθε σύμβολο κατηγορήματος p και μεταβλητές $X_1, X_2, \ldots, X_n$, $p(X_1, X_2, \ldots, X_n)$ είναι μια ατομική πρόταση
  2. αν X, Y είναι μεταβλητές και $\theta \in \{<, >, =, \leq, \geq, \neq\}$, τότε $X\ \theta\ Y$ είναι μια ατομική πρόταση

  Σημείωση: ένα σύμβολο κατηγορήματος p δηλώνει μια σχέση με την ερμηνεία ότι $p(X_1, \ldots, X_n)$ είναι αληθές αν $(X_1, \ldots, X_n)$ ανήκει στη σχέση που αντιστοιχεί στο p

- Αν οι F_1, F_2 είναι καλα-δομημένες προτάσεις, τότε και οι $F_1 \wedge F_2, F_1 \vee F_2$ και $\neg F_1$ είναι καλά-δομημένες προτάσεις.
- Αν η μεταβλητή X είναι ελεύθερη στην πρόταση F, τότε οι $(\exists X)F$ και $(\forall X)F$ είναι καλά-δομημένες προτάσεις.

Σημείωση: η χρήση ενός ποσοδείκτη $(\exists, \forall)$ δεσμεύει κάθε στιγμιότυπο μιας μεταβλητής μέσα σε μια πρόταση

Προτεραιότητα Τελεστών:

1. $\neg, \exists, \forall$ – από δεξιά προς τα αριστερά
2. $\wedge$ – από αριστερά προς τα δεξιά
3. $\vee$ – από αριστερά προς τα δεξιά

Παράδειγμα: η πρόταση

$$(\forall X)\neg p(X,Y) \vee q(Y) \wedge r(X)$$

ομαδοποιείται ως

$$((\forall X)(\neg p(X,Y))) \vee (q(y) \wedge r(X))$$

Σχεσιακός Λογισμός Πεδίων (DRC)

- κάθε πρόταση του ΣΛΠ με μία ή περισσότερες ελεύθερες μεταβλητές ορίζει μία σχέση της οποίας τα γνωρίσματα αντιστοιχούν στις ελεύθερες μεταβλητές.
- Οι μεταβλητές συνιστούν πλειάδες της σχέσης και λαμβάνουν τιμές από τα αντίστοιχα πεδία τιμών.
- Μια επερώτηση στη γλώσσα του ΣΛΠ έχει τη μορφή: $\{X_1, \ldots, X_n | F(X_1, \ldots, X_n)\}$, όπου η F είναι μια καλά-δομημένη πρόταση του ΣΛ με ελεύθερες μεταβλητές $X_1, \ldots, X_n$
- Η έκφραση $\{X_1, \ldots, X_n | F(X_1, \ldots, X_n)\}$ αναπαριστά το σύνολο των πλειάδων $(a_1, \ldots, a_n)$ οι οποίες είναι τέτοιες ώστε, αν η τιμή a_i αντικατασταθεί για τη μεταβλητή X_i, για όλα τα $1 \leq i \leq n$, τότε η πρόταση $F(a_1, \ldots, a_n)$ είναι αληθής.

- Σχέσεις που ορίζονται απο εκφράσεις του ΣΛΠ μπορεί να είναι μη-πεπερασμένες.

  Π.χ., η έκφραση $\{X, Y | \neg p(X, Y)\}$ ορίζει μια μη-πεπερασμένη σχέση.

- Ενα υποσύνολο των προτάσεων του ΣΛΠ, οι **ασφαλείς** προτάσεις, ορίζουν πεπερασμένες σχέσεις.

Παραδείγματα

- Find the ids of customers that live in New York.

  $\{C \mid (\exists N)(\exists T)(\exists D)\ Customers(C, N, T, D)\ \wedge\ (T = "NewYork")\}$

- Find the names and unit prices of products ordered by customer c002 through agent a01.

  $\{N, P \mid (\exists I)(\exists C)(\exists Q)(\exists O)(\exists M)(\exists QT)(\exists D)\ Products(I, N, C, Q, P)\ \wedge\ Orders(O, M, c001, a01, I, QT, D)\}$

Από τη Σ.Α. στο Σ.Λ.Π

Θεώρημα: Κάθε έκφραση της σχεσιακής άλγεβρας μπορεί να εκφραστεί στο Σχεσιακό Λογισμό Πεδίων.

Απόδειξη: Θα δείξομε, με επαγωγή στον αριθμό των τελεστών της αλγεβρικής έκφρασης, ότι για κάθε έκφραση E της Σ.Α. η οποία ορίζει μια σχέση βαθμού k, υπάρχει μια πρόταση $F(X_1, \ldots, X_k)$ του ΣΛΠ η οποία ορίζει την ίδια σχέση.

Βάση επαγωγής: $E = R$, για κάποια σχέση R. Η αντίστοιχη πρόταση του ΣΛΠ είναι $r(X_1, \ldots, X_k)$.

Υπόθεση επαγωγής: έστω ότι για τις εκφράσεις E_1, E_2 της Σ.Α. οι οποίες ορίζουν σχέσεις βαθμού k υπάρχουν προτάσεις F_1, F_2 του ΣΛΠ οι οποίες ορίζουν τις ίδιες σχέσεις

Επαγωγικό βήμα: Θα δείξομε ότι υπάρχουν προτάσεις του ΣΛΠ οι οποίες ορίζουν τις ίδιες σχέσεις με τις εκφράσεις $E_1 \cup E_2, E_1 - E_2, \pi_X(E_1), E_1 \times E_2$ και $\sigma_F(E_1)$.

1. $E = E_1 \cup E_2$ και έστω k ο βαθμός των σχέσεων των $\mathrm{E}, \mathrm{E}_1, \mathrm{E}_2$. Από την επαγωγική υπόθεση έχουμε ότι υπάρχουν προτάσεις F_1, F_2 οι οποίες ορίζουν τις ίδιες σχέσεις με τις E_1, E_2. Υποθέτομε επίσης ότι οι F_1, F_2 έχουν τις ίδιες ελεύθερες μεταβλητές $X_1, \ldots, X_k$ (αν όχι, μπορούμε να μετονομάσομε τις μεταβλητές). Οι μεταβλητές $\mathrm{X}_1, \ldots, \mathrm{X}_\kappa$ αντιστοιχούν στα γνωρίσματα των σχέσεων E, E_1, E_2. Τότε, η πρόταση που αντιστοιχεί στην έκφραση E είναι η πρόταση $F_1 \vee F_2$.

2. $E = E_1 - E_2$ και έστω ότι οι προτάσεις $F_1(X_1, \ldots, X_k)$ και $F_2(X_1, \ldots, X_k)$ αντιστοιχούν στις εκφράσεις E_1 και E_2 αντίστοιχα. Τότε, η πρόταση που αντιστοιχεί στην Ε είναι η $F_1 \wedge \neg F_2$.

3. $E = \pi_{A_{i_1},\ldots,A_{i_n}}(E_1)$. Εστω $\{A_1,\ldots,A_k\}$ τα γνωρίσματα της σχέσης της E_1. Εστω ακόμη $\{A_{i_1},\ldots,A_{i_n}\}$ και $\{A_{j_1},\ldots,A_{j_m}\}$ υποσύνολα του $\{A_1,\ldots,A_k\}$ έτσι ώστε $\{A_{i_1},\ldots,A_{i_n}\} \cap \{A_{j_1},\ldots,A_{j_m}\} = \emptyset$ και $\{A_{i_1},\ldots,A_{i_n}\} \cup \{A_{j_1},\ldots,A_{j_m}\} = \{A_1,\ldots,A_k\}$. Μια πλειάδα t ανήκει στη σχέση της E αν και μόνο αν υπάρχουν τιμές για τα γνωρίσματα $A_{j_1},\ldots,A_{j_m}$ με τις οποίες μπορούμε να "επεκτείνομε" την t για να πάρομε μια πλειάδα της E_1. Η πρόταση για την Ε είναι $F(X_{i_1},\ldots,X_{i_n}) = (\exists X_{j_1})(\exists X_{j_2})\ldots(\exists X_{j_m})\ F_1(X_1,\ldots,X_k)$

4. $E = E_1 \times E_2$. Εστω $F_1(X_1,\ldots,X_n)$ και $F_2(Y_1,\ldots,Y_m)$ προτάσεις του ΣΛΠ για τις E_1 και E_2 αντίστοιχα. Υποθέτουμε ότι $\{X_1,\ldots,X_n\} \cap \{Y_1,\ldots,Y_m\} = \emptyset$. Τότε η πρόταση για την E είναι $F(X_1,\ldots,X_n,Y_1,\ldots,Y_m) = F_1(X_1,\ldots,X_n)\ \wedge\ F_2(Y_1,\ldots,Y_m)$

5. $E = \sigma_{A_i\theta A_j}(E_1)$. Εστω $F_1(X_1,\ldots,X_n)$ η πρόταση του ΣΛΠ για την Ε$_1$. Τότε η πρόταση για την E είναι $F_1\ \wedge\ (X_i\theta X_j)$

Παράδειγμα

$$E = R(X,Y) - (S(X) \times \pi_Y(R(X,Y))).$$

Η πρόταση του ΣΛΠ για την E είναι η:

$$F(X,Y) = r(X,Y) \ \wedge \ \neg(s(X) \ \wedge \ (\exists Z) \ r(Z,Y))$$

Προτάσεις Ανεξάρτητες Πεδίου

- Καθε μεταβλητή σε μια πρόταση του ΣΛΠ παίρνει τιμές από καποιο πεδίο τιμών. Αυτό το πεδίο τιμών περιλαμβάνει τις τιμές οι οποίες εμφανίζονται στην ίδια την πρόταση και στη βάση δεδομένων. Για μια πρόταση F, το σύνολο των πεδίων τιμών όλων των μεταβλητών της συμβολίζεται με $DOM(F)$.
- Το σύνολο $DOM(F)$ δεν είναι σταθερό για μια δεδομένη πρόταση, αλλά εξαρτάται από τις σχέσεις που αντιστοιχούν σε κάθε κατηγόρημα της πρότασης.
- Το $DOM(F)$ μπορεί να εκφραστεί ως η ένωση των συνόλων των σταθερών που εμφανίζονται στην F και τις προβολές όλων των σχέσεων που αντιστοιχούν σε κατηγορήματα της F σε κάθε τους γνώρισμα.

Παράδειγμα

Εστω P, Q οι σχέσεις που αντιστοιχούν στα κατηγορήματα p, q αντίστοιχα. Τότε, για την πρόταση

$$F = p(X, Y) \wedge q(Y, Z) \vee X > 10$$

$$DOM(F) = \{10\} \cup \pi_X(P) \cup \pi_Y(P) \cup \pi_Y(Q) \cup \pi_Z(Q)$$

Διαισθητικά, μια πρόταση λέγεται **ανεξάρτητη πεδίου** αν η σχέση την οποία ορίζει δεν μπορεί να περιέχει πλειάδες οι οποίες περιλαμβάνουν σταθερές που δεν ανήκουν στο πεδίο της.

Ορισμός: Εστω $F(X_1, \ldots, X_n)$ μια πρόταση του ΣΛΠ και $D \supseteq DOM(F)$ ένα σύνολο τιμών. Η σχέση της F αναφορικά με το D είναι το σύνολο των πλειάδων $(a_1, \ldots, a_n)$ του D^n, οι οποίες είναι τέτοιες ώστε, όταν κάθε a_i αντικατασταθεί για το X_i, η $F(a_1, \ldots, a_n)$ είναι αληθής. Η F λέγεται ανεξάρτητη πεδίου (domain independent) αν η σχέση της αναφορικά με το D δεν εξαρτάται από το ίδιο το D. Αν η F είναι ανεξάρτητη πεδίου, τότε η σχέση της αναφορικά με οποιοδήποτε σύνολο D είναι η ίδια με τη σχέση της με το σύνολο $DOM(F)$.

Παραδείγματα

1. Η πρόταση $F_1 = \neg r(X, Y)$ δεν είναι ανεξάρτητη πεδίου.

   Εστω R η σχέση για το κατηγόρημα r.
   $DOM(F_1) = \pi_X(R) \cup \pi_Y(R)$. Αν D είναι οποιοδήποτε σύνολο που περιέχει το $DOM(F_1)$, τότε η σχέση για την F_1 αναφορικά με το D είναι $D \times D - R$. Αν a ανήκει στο D αλλά όχι στο $DOM(F_1)$, τότε η πλειάδα (a, a) ανήκει στη σχεση της F_1 αναφορικά με το D, αλλά όχι στη σχέση της F_1 αναφορικά με το $DOM(F_1)$.

2. $F_2 = (\exists Y)(p(X, Y) \ \vee \ q(Y, Z))$

   Εστω $P = \{(a, b), (c, d)\}$ και $Q = \{(e, f)\}$ οι σχέσεις για τα κατηγορήματα p και q αντίστοιχα. Τότε
   $DOM(F_2) = \{a, b, c, d, e, f\}$. Εστω $D = \{a, b, c, d, e, f, g\}$. Τότε η σχέση της F_2 αναφορικά με το D περιλαμβάνει την πλειάδα (a, g).

Η σχέση της F_2 θα είναι διαφορετική αν το στοιχείο g αφαιρεθεί από το D. Αρα, η F_2 δεν είναι ανεξάρτητη πεδίου.

3. $F_3 = (\exists Y)(p(X,Y) \ \wedge \ q(Y,Z))$

   Εστω (a,b) μια πλειάδα στην σχέση της F_3 αναφορικά με ένα σύνολο D. Τότε, πρέπει να υπάρχει μια τιμή c τέτοια ώστε $p(a,c) \ \wedge \ q(c,b)$ να είναι αληθής. Αρα, (a,c) ανήκει στην P και (c,b) ανήκει στην Q. Επομένως, a ανήκει στην $\pi_X(P)$, b ανήκει στην $\pi_Z(Q)$ και a και b ανήκουν στο $DOM(F_3)$. Για οποιοδήποτε σύνολο $D \supset DOM(F_3)$, το σύνολο των ζευγών τιμών (a,b) στη σχέση της F_3 αναφορικά με το D θα είναι το ίδιο με τη σχέση της F_3 αναφορικά με το $DOM(F_3)$. Αρα, η F_3 είναι ανεξάρτητη πεδίου.

Ασφαλείς Προτάσεις

- Η ανεξαρτησία πεδίου είναι μια έννοια η οποία αφορά τη σημασιολογία των προτάσεων του ΣΛ. Δεν υπάρχει αλγόριθμος ο οποίος να αποφασίζει αν μια πρόταση είναι ανεξάρτητη πεδίου.
- Το σύνολο των **ασφαλών** προτάσεων του ΣΛ είναι ένα υποσύνολο του συνόλου των προτάσεων που είναι ανεξάρτητες πεδίου.
- Υπάρχει αλγόριθμος ο οποίος βασίζεται στη σύνταξη και ο οποίος αποφασίζει αν μια πρόταση είναι ασφαλής ή όχι.

Ορισμός: Μια πρόταση του ΣΛΠ είναι ασφαλής αν ικανοποιούνται οι παρακάτω ιδιότητες:

1. Δεν χρησιμοποιείται ο καθολικός ποσοδείκτης ($\forall$).
2. Αν $F = F_1 \vee F_2$ τότε F_1 και F_2 έχουν τις ίδιες ελεύθερες μεταβήτές.
3. Για οποιαδήποτε μέγιστη υποπροταση $F_1 \wedge \ldots \wedge F_m$ της F όλες οι ελεύθερες μεταβλητές των F_i πρέπει να είναι περιορισμένες με την εξής έννοια:
   (a) Μια μεταβλητή είναι περιορισμένη αν είναι ελέυθερη σε μια υποπρόταση F_i, όπου η F_i δεν είναι μια αριθμητική σύγκριση και δεν προηγείται $\neg$.
   (b) Αν F_i είναι $X = a$ όπου a είναι σταθερά, τότε η X είναι περιορισμένη.

(c) Αν F_i είναι $X = Y$ και η Y είναι περιορισμένη, τότε η X είναι περιορισμένη.

4. Αν η F είναι της μορφής $H_1 \wedge \ldots H_j \wedge \neg G \wedge I_1 \wedge \ldots I_j$ και οι συζεύξεις ικανοποιούν τον παραπάνω κανόνα, τότε είναι ασφαλής αν τουλάχιστον ένα από τα H ή I δεν έχει άρνηση.

Παραδείγματα

1. Η $X = Y$ δεν είναι ασφαλής καθώς καμία από τις μεταβήτές X, Y δεν είναι περιορισμένη.
2. Η $(X = Y) \wedge P(X, Y)$ είναι ασφαλής.
3. Η $(X = Y) \vee P(X, Y)$ δεν είναι ασφαλής.
4. Η $r(X, Y, Z) \wedge \neg(p(X, Y) \vee (q(Y, Z))$ δεν είναι ασφαλής. Μπορεί όμως να μετατραπεί σε μια ασφαλή προταση με χρήση του κανόνα De Morgan. Η ισοδύναμη ασφαλής πρόταση είναι:
   $r(X, Y, Z) \wedge \neg p(X, Y) \wedge \neg q(Y, Z)$.