FIATAL MŰSZAKIAK TUDOMÁNYOS ÜLÉSSZAKA

Kolozsvár, 1997, március 21-23

ASSZOCIATIVITÁS SZEREPE AZ ADATBÁZIS LEKÉRDEZŐ FELÜLETEKNÉL

Kovács László

Bevezetés

Az adatbázis kezelő rendszerek (DBMS) fejlődése során már a kezdetek óta megjelent az igény a rugalmas, emberközelű kezelő felület kialakítására. A hagyományos, relációs adatmodellek által igényelt megkötések felszabadítására számos kutatási projekt futott és fut ma is, melyek döntő többsége a relációs modell egyfajta módosításán alapulnak. A relációs modellre épülő javaslatok mellett újszerű, elsődlegesen szemantikai alapú adatmodellek is megszülettek a hagyományos adatkezelés rugalmasságának fokozására. A hagyományos kezelő felületeknek a pozitív vonásaik, mint pl. a hatékony működés, mellett számos olyan oldaluk is van, melyekben van még lehetőség a továbbbővítésre.

A hagyományos, pl SQL alapú kezelő felületek alkalmazása számos előfeltétel teljesülését követeli meg a felhasználótól, ami a rendszer rugalmas használhatóságát jelentősen csökkenti. Bosc és Piver[1] az alábbi lényeges megkötésekre hívja fel a figyelmet:

- a felhasználónak ismernie kell az adatmodellt,
- ismernie kell az adatbázis tartalmát,
- ismernie kell a lekérdező nyelv szintaktikáját,
- meghatározott lekérdezési céllal kell rendelkeznie.

Az egyes megkötések feloldására különböző megoldások születtek, melyek más-más kutatási területre visznek el minket: szintaktikai hibák automatikus kijavítása, szemantikai hibák korrigálása, lehetséges válaszok kiértékelése, fuzzy fogalmak bevonása, szabad mozgás az adatbázisban. A továbbiakban ez utóbbi területtel foglalkozunk részletesebben.

A vizsgált dokumentumban történő szabadabb mozgások jellemzője, hogy a felhasználó rugalmasabban lépegethet az egyes elemek között, a szomszédos elemeket összekötő útvonalak mellett lehetősége van nagyobb ugrások elvégzésére is. Természetesen az ugrások rendszerint nem véletlenszerűek, hanem valamilyen logikai vonal mentén történnek. E mozgások alapja a fogalmak

közötti társítás, az asszociáció. A különböző fogalmak között különböző erősségű kapcsolatok lehetségesek.

Az adatbázis kezelés területén az adatmodellek struktúrális vagy integritási részének egyik fő eleme az egyedek közötti kapcsolatok leírására szolgáló rész. A relációs adatmodellben például a kulcs-kapcsolókulcs (idegenkulcs) páros szolgál a kapcsolatok leírására. Az adatbázisban tárolt egyedek között tehát ezen modellelem révén teremthetjük meg a kapcsolatot, az asszociációt. A fejlesztések célja, olyan kezelő felületek, adatmodellek kidolgozása, melyek minnél természetesebben, minnél rugalmasabban teszik az asszociációt elérhetővé a felhasználó számára.

Az eddigi kutatások áttekintése

A 1980-as évek elején jött létre a relációs adatmodell módosításaként az úgynevezett Universal Relation [2] adatmodell. A modellnek több változata is van, melyből a legkidolgozottabb az ún. 'weak universal model'. Ebben a modellben az összes attributumot tartalmazó és az egész adatbázist reprezentáló univerzális reláció csak a felhasználó modellképében létezik. Az adatok továbbra is normál relációkban tárolódnak. A modell főbb struktúrális elemei:

- attributum,
- objektum (attributumok közötti függőségi kapcsolat),
- maximális objektum (a társítható attributumok körét szabja meg).

A lekérdezésekben az attributum nevekre és a virtuális rekord azonosítókra kell hivatkozni. A modell egyik fő erőssége a rugalmasság növelésében, hogy a felhasználó a kérdéseit fogalom, vagyis attributum orientáltan adhatja meg. A rendszer bizonyos határok között maga határozza meg, hogy a lekérdezés végrehajtásához mely kapcsolatokat használja fel. Mivel nincsenek különálló relációk a felhasználó nagyrészt tetszőlegesen társíthatja az egyes attributumokat a lekérdezés megfogalmazása során.

A relációs modellből kiinduló javaslatok mellett számos szemantikai adatmodell (SDM) is létrejött az adatkezelés rugalmasságának fokozására. Az SDM rendszerekben az adatkezelés az emberi gondolkodáshoz közel álló objektumok szintjén valósul meg [6]. Számos SDM modellhez készítettek emberközelű lekérdező felületet. Wong és Kuo [4] az ER modelhez, míg Heiler és Rosenthal [5] pedig a funkcionális adatmodellel rendelkező DAPLEX-hez készített lekérdező rendszert. A funkcionális adatmodellek fő jellemzője, hogy az egyes egyedek közötti kapcsolat a függvény mechanizmussal valósul meg, s a függvények egymásba ágyazásával, egymásutáni meghívásukkal sokkal rugalmasabban lehet mozogni, mint a hagyományos kulcs - idegen kulcs szerkezettel.

Az SDM modellek igen közeli kapcsolatban állnak az objektum orientált elveken nyugvó objektum orientált adatmodellekkel (OODM). A objektumok között megvalósuló összetettebb kapcsolatok tágabb lehetőséget biztosítanak az egyedek közötti kapcsolatok, asszociációk megvalósításának is. E kapcsolatokon nyugvó adatbázis lekérdező rendszer kifejlesztése igen aktuális kutatási problémakör. Egy

grafikus, windows technikán alapuló rendszer a PESTO[3]. A rendszerben külön ablak tartozik minden objektum típushoz. Az objektumok közötti kapcsolatok előbb egy hivatkozás jelzővel jelenik meg, melyet a felhasználó kifejthet, megkapva a hivatkozott objektumot leíró ablakot. A fő objektumtípus egyed típus váltásánál a többi típus egyedei is szinkronban mozognak. E grafikus lekérdező felület előnye, hogy a felhasználó láthatja, hogy mely irányokban lehet továbblépni. Ezért nincs szükség az egyes objektumok pontos szerkezetének ismeretére, elegendő csupán rámutatni arra a kapcsolatra, mely irányban tovább szeretnénk lépni.

AZ SFQ adatmodell

A rugalmas és az asszociatívuson alapuló adakezelő felület vizsgálatához kidolgoztunk egy saját SDM adatmodellt, melyre SFQ néven hivatkozunk. E adatmodell fő vonásaiban a funkcionális adatmodellhez áll közel. A modell legfontosabb strukturális jellemzői:

- az egyedek lehetnek atomiak és összetettek,
- az egyed típusok nem szerepelnek külön modellegyedként,
- csak az atomi egyedek értékei érhetők el,
- három kapcsolattípus definiált: normál, IS_A, kulcs kapcsolat,
- minden összetett egyedhez léteznie kell kulcs kapcsolatrendszernek.

A modell grafikus elemeit mutatja be a következő ábra.

1. ábra: SFQ strukturális elemei

A modell műveleti részében a relációs algebrához hasonló halmazorientált szemlélettel dolgozik. A lekérdezések operandusai egyedek halmazai és eredményük is egyedhalmaz. A következő műveleteket tartalmazza az SDM adatmodell:

- alapszelekció
- Descartes szorzat
- kollekció
- asszociáció

Az alapszelekció egy egyedfajta előfordulás halmazából a megadott feltételnek elget tévő előfordulásokat adja vissza:

$$\sigma\{D \mid \text{feltétel}\}$$

A Descartes szorzat a hagyományos join művelethez hasonlóan több egyed típus előfordulásainak párosításából választja ki a megfelelő párokat. A művelet formális leírása:

$$\sigma\{D_1 \times D_2 \times \dots \times D_n \mid \text{feltétel}\}$$

A kollekció esetén egy fő egyedtypus egyazon előfordulásához kapcsolódó más egyedtypusok értékei alkotják a megoldás eredményét. Ezzel a művelettel rugalmasan hozhatunk létre saját egyedi táblákat, nézeteket. A kapcsolódás útvonala az adatbázis sémából adódik. Alapesetben az egyes típusok közötti legrövidebb láncolat kerül meghatározásra. Más irányú kapcsolódási útvonal kijelöléséhez a felhasználónak érintési pontokat, egyedeket kell előírnia. A kollekció formátuma:

$$\sigma\{D1 \rightarrow (D2, \dots, Dn) \mid \text{feltétel}\}$$

Az asszociáció nem az egyedtypusok közötti útvonal meghatározásán alapszik, hanem az egyes egyedelőfordulások közötti kapcsolódáson. Két egyedelőfordulás asszociációban áll, ha létezik az adatbázis előfordulásban az egyiktől a másikig vezető út. A művelet jele:

$$\sigma\{D1 \circ D2 \mid \text{feltétel}\}$$

A lekérdezésekben az egyedtypusok és az egyedelőfordulásaik között tetszőleges, rugalmasan alakítható kapcsolatokat állíthatunk fel. A modellben például a fiatal gépkocsi-tulajdonosok gyerekeinek nevei az alábbi utasítással kérdezhető le:

$$\sigma\{\text{autó} \rightarrow \text{tulajdonos} \rightarrow \text{gyerek} \rightarrow \text{név} \mid \text{fiatal}(\text{tulajdonos} \rightarrow \text{kor})\}$$

A modell jelenlegi fázisában a modell implementálásán dolgozunk. Az elkészítendő rendszer kisebb egyfelhasználós rendszer kezelésére lesz alkalmas. Dolgozunk az SDM modell és a relációs adatmodell konverziós kapcsolatának a kiépítésén is. Jelenleg egy inverz normalizációs folyamat alkotja a konverzió magját. Az inverz normalizáció segítségével létező relációs adatmodellből a tervező szemantikai információ alapján több lépésen keresztül előáll az SDM modell. Az így előálló konverziós szabályok segítségével előállítható a relációs adatbázisból az SDM adatbázis.

Irodalomjegyzék

- [1] P. Bosc - O. Pivert, Fuzzy Querying in Conventional Databases, Fuzzy Logic for Management of Uncertainty, 1993
- [2] D. Maier - J. Ullman - M. Vardi, On the Foundation of the Universal Relation Model, ACM TODS 9/2, 1984
- [3] M. Carey - L. Haas - V. Maganty - J. Williams, PESTO: An Integrated Query/Browser for Object Databases, VLDB96, 1996
- [4] Wong-Kuo, GUIDE, A Graphical User Interface for Database Exploration, VLDB82, 1982
- [5] Heiler-Rosenthal, A Visual Interface for the Functional Model with Recursion, VLDB85, 1985
- [6] Lyngbeak - Kent, A Data Modeling Methodology for the Design and Implementation of Information Systems, On OODBMS, 1991