

Ondersteuning van identiteit maakt weg vrij voor nieuwe relationele concepten

# 'The Third Manifesto' kritisch bekeken (2)

Maurice Gittens

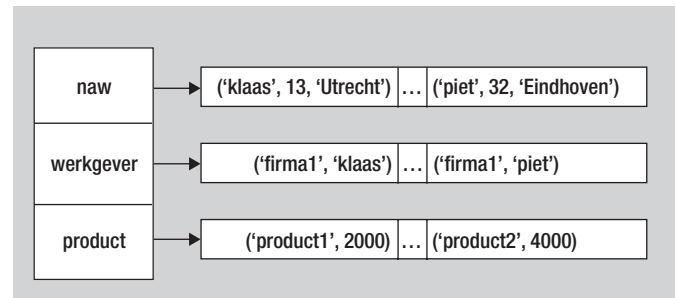
**In 'Foundation for future database systems; The Third Manifesto' (TM) van Darwen en Date is geen plaats ingeruimd voor identiteit in toekomstige databases. De auteurs categoriseren de ondersteuning van (object) identiteit, in tegenstelling tot E.F. Codd, als een blunder. Dit artikel zal een informele motivatie voor ondersteuning van objectidentiteit in toekomstige databases presenteren aan de hand van door de predikatenlogica geïnspireerde concepten als identiteitsdomeinen en identiteitsoperatoren.**

In het boek 'Foundation for future database systems; The Third Manifesto' wordt door Date en Darwen objectidentiteit verworpen. In dit artikel wordt in vogelvlucht een schets gegeven van mogelijkheden die in toekomstige databases door de invalshoek van Date en Darwen worden uitgesloten. Om het denkkader van waaruit dit artikel geschreven is, inzichtelijk te maken wordt in dit artikel eerst ingegaan op de logische fundamenten van relationele systemen. Vervolgens worden uitbreidingsvoorstellen voor relationele systemen gedaan die analoog zijn aan bestaande uitbreidingen van de eerste orde predikatenlogica. De uitbreidingsvoorstellen worden met voorbeelden in een SQL-achtige taal inzichtelijk gemaakt.

## Logische interpretatiefuncties

Een wijze waarop een appreciatie voor identiteit in toekomstige databases ontwikkeld kan worden is door terug te grijpen op het fundament van relationele systemen; namelijk de predikatenlogica. In de predikatenlogica wordt informatie over een bepaald domein afgebeeld op een abstractie die bekend staat als een interpretatiefunctie – volgens Van Eijck en Thijsse (zie de literatuurreferentie) is dit hulpmiddel voor het eerst door de Poolse logicus

Maurice Gittens werpt een kritische blik op het boek 'Foundation for future database systems; The Third Manifesto' van Hugh Darwen en Chris Date. Na de twee vermeende grote blunders te hebben besproken in DB/M 5 bespreekt hij in dit artikel onder andere identiteitsdomeinen en identiteitsoperatoren. In een derde vervolgartikel zullen Chris Date en Hugh Darwen reageren op de stellingen van Gittens.



**Afbeelding 1:** Vereenvoudigde interpretatiefunctie uit de predikatenlogica.

Alfred Tarski ingezet bij het expliciet formuleren van de semantiek van de predikatenlogica in 1933.

Men zou kunnen zeggen dat een interpretatiefunctie de logische equivalent van een relationele database is. In afbeelding 1 zijn drie relaties weergegeven in een vereenvoudigde interpretatiefunctie uit de predikatenlogica.

Domein van de interpretatiefunctie
naw
werkgever
product

Alle relationele databases kunnen in principe als vereenvoudigde logische interpretatiefuncties worden gezien. In dit artikel wordt in zo'n geval gesproken van de logische representatie van de relationele database. De bovenstaande schematische voorstelling van een interpretatiefunctie staat de onderscheiding van een aantal wezenlijke concepten toe. Een aantal hiervan wordt in het kort toegelicht.

## Het domein van de interpretatiefunctie

Het eerste concept dat in deze onderscheiden wordt is het *domein* van de interpretatiefunctie (de term domein wordt hier in wiskundige zin gebruikt: "Het domein en het bereik van een wiskundige functie"). Dat is in het getoonde voorbeeld de verzameling relatie namen in de eenvoudige database. De naam (of preciezer gezegd de identiteit) van een predikaat (of relatie als men dat liever heeft) wordt door de interpretatiefunctie geassocieerd met de extensie van de relatie (dat is de verzameling tuples waaruit de relatie bestaat).

## Het bereik van de interpretatiefunctie

Het *bereik* van een logische interpretatiefunctie is natuurlijk op zijn beurt de verzameling extensies van alle relaties in de voorbeeld-database. In dit verband kan dus de *extensie* van een relatie onderscheiden worden van de *identiteit* van de relatie. Het bereik van de interpretatiefunctie in het voorbeeld wordt weergegeven zoals hieronder.

Bereik van de interpretatiefunctie
{('klaas', 13, 'Utrecht'), ..., ('piet', 32, 'Eindhoven')}
{('firma1', 'klaas'), ..., ('firma1', 'piet')}
{('product1', 2000), ..., ('product2', 4000)}

In het volgende tabel wordt de extensie van de productrelatie uit de voorbeeld-database getoond.

Extensie van de productrelatie
('product1', 2000)
...
('product2', 4000)

In termen van de TM uitgedrukt, geldt dat een relatiewaarde (relation value) de extensie van een relatie is, terwijl een relatievariabele (relation variable of relvar) te zien is als de koppel gevormd door de identiteit van een relatie en de extensie van dezelfde relatie. De TM kiest om op vrij dogmatische wijze voorbij te gaan aan de identiteit van een relatie.

## Voor wiskundige tuples geldt dat de volgorde van de tuple-objecten wel van belang is

Tussen logische systemen en relationele systemen bestaan er overeenkomsten en natuurlijk ook de nodige verschillen. Bijvoorbeeld: voor wiskundige tuples geldt dat de volgorde van de tuple-objecten wel van belang is. Voor relationele systemen is dat juist niet het geval. Mede als gevolg van dit feit is bijvoorbeeld in relationele systemen de volgorde van kolommen in een tabel niet van belang. In de predikatenlogica juist wel. In deze beschouwing worden zulke verschillen als cosmetische verschillen gezien, die voor de logische uitdrukingskracht van de betreffende formalismen niet relevant zijn.

## Identiteitsdomeinen en Identiteitsoperatoren

De relationele algebra is in essentie een verzameling operatoren die *relatiewaarden* manipuleren en transformeren. In dit artikel wordt concreet voorgesteld om toekomstige relationele systemen uit te breiden met de ondersteuning van identiteitsdomeinen en identiteitsoperatoren.

- Identiteitsoperatoren: dus het wordt mogelijk om te itereren over het *domein* van de logische interpretatiefunctie waardoor relationele abstracties zoals tabellen en domeinen toegankelijk worden (geïnspireerd door de tweede orde predikatenlogica);
- Identiteitsdomeinen: dus het wordt mogelijk om de identiteit van relationele abstracties (zoals tabellen en domeinen) als attribuut waarde in tabellen op te slaan (analoog aan de introductie van zogenaamde predikaatconstanten in de extensionele typenlogica).

Voor de hand liggende identiteitsoperatoren die toekomstige databases analoog aan de predikatenlogica zouden kunnen inzetten, zijn de *foreach* en ook bijvoorbeeld de *exists* operator. Aan de hand van een aantal voorbeelden, uitgedrukt in een SQL-achtige query-taal, wordt informeel de *foreach* operator geïntroduceerd.

Eerste voorbeeld:

```
foreach relation r select * from r;
```

Dit statement selecteert met behulp van de *foreach* operator alle tuples van een database (we doen hier niet moeilijk over schema's etcetera). Het resultaat van deze query zou prima in een XML-representatie gegoten kunnen worden. Het is belangrijk om hierbij op te merken dat:

- De variabele *r* gebonden is aan de identiteit en niet aan de waarde van de relatie;
- Het type van *r* in het bovenstaande een relatie-type is en niet bijvoorbeeld van het type string is;
- de *foreach* operator iteratie over het *domein* van logische interpretatiefuncties toe laat en *niet* over de extensie van individuele relatievariabelen in een database;
- het selecteren van alle tuples in een relationele database op basis van bestaande relationele operatoren niet mogelijk is.

Nog enkele voorbeelden:

```
foreach relation r where r.owner = "maurice"
    select * from r;
foreach relation r foreach attribute a in r select
    a.name where a.domain = integer;
```

## Identiteitsdomeinen

Eenvoudige voorbeelden van het gebruik van identiteitsdomeinen zijn:

```
select tabelnaam from eenTabel as naam select *
    from naam;
select (select domeinnaam1 from eenTabel, select
    domeinnaam2 from eenTabel) from voorbeeldTabel;
```

Deze voorbeelden tonen hoe de identiteit van een tabel uit een tabel gelezen zou kunnen worden om vervolgens de inhoud van de betreffende tabel te selecteren, en hoe de identiteit van twee attributen uit een tabel gelezen zou kunnen worden om vervolgens als kolomfilters te worden ingezet.

---

## Conclusies

Identiteitsdomeinen en identiteitsoperatoren zijn voorgesteld als concepten die in de ene of andere vorm in toekomstige databases dienst zouden kunnen doen. Identiteitsdomeinen en identiteitsoperatoren zijn een exponent van de overtuiging dat het in de discussie over toekomstige databases niet zozeer zou moeten gaan

## Men zou kunnen zeggen dat een interpretatiefunctie de logische equivalent van een relationele database is

over relationeel versus objectgeoriënteerd, maar veeleer over het orthogonaliseren van de primitieve logische concepten die ten grondslag liggen aan relationele systemen en ook objectgeoriënteerde systemen, met als doel een coherent geheel te vormen.

Over identiteit in toekomstige relationele databases is natuurlijk veel meer te zeggen dan in deze korte informele introductie is gezegd. De ondersteuning van identiteit in toekomstige databases maakt de weg vrij voor de inzet van nieuwe relationele concepten als orthogonalen, partitioneringen, vlakken etcetera, die de kwaliteit en beheersbaarheid van toekomstige operationele databases en management-informatiesystemen ten goede zal kunnen komen. Mede op basis van het in dit artikel gepresenteerde materiaal wordt tot slot geconcludeerd dat de verwerping van identiteit in toekomstige databases door Date en Darwen niet een dictaat van de prudentie is.

### Literatuur

*C.J. Date, Hoge Darwen (2000), Foundation for Future Database Systems. Addison-Wesley Publishing Company.*

*E.F. Codd, (1979), Extending the Database Relational Model to Capture more Meaning.*

*Jan van Eijck en Elias Thijsse (1989), Logica voor alfa's en informatici. Academic Service.*

**Maurice Gittens** (maurice@gittens.nl) is zelfstandig IT-consultant.