

The Third Manifesto: Setting the Record Straight (4)

Nulls en driewaarden logica

H. Darwen en C. J. Date

In The Third Manifesto verwerpen de auteurs het gebruik van onbekende waarden zoals NULL en de toepassing van logica met meer dan twee waarden (WAAR en NIET WAAR). Maar onderbouwen ze die afwijzing eigenlijk wel goed, zo vroeg Maurice Gittens zich af, en staat dat alles niet haaks op de ideeën van Ted Codd? In dit artikel gaan Date en Darwen diep in op deze vraag.

De paragraaf in Gittens' artikel over dit onderwerp [8] heet: Geen grondige onderbouwing voor de afwijzing van onbekende waarden. Dit verwijst impliciet naar RM Proscriptie 4 ('NO NULLS') in The Third Manifesto [7] kortweg TTM¹, waarin staat: *D shall include no concept of a 'relation' in which some 'tuple' includes some attribute that does not have a value.*

The naam D refereert in het algemeen aan elke computertaal die zich conformeert aan de proscripties, de voorschriften van The Third Manifesto. De proscripties, de verboden, staan alleen maar in TTM om reden van verduidelijking en verdere uitwerking – ze volgen allemaal logisch uit de voorschriften. De meeste verboden zijn opgenomen juist als SQL in strijd is met het betreffende voorschrift, zoals hier het geval is.

In DB/M 2 van 2007 uitte Maurice Gittens nogal wat kritiek op de derde editie van het standaardwerk 'Databases, Types, and the Relational Model: The Third Manifesto' van Hugh Darwen en Chris Date. Gittens vraagt zich af wat de bijdrage van Darwen en Date aan het relationele model de afgelopen jaren is geweest – de derde editie van TTM acht hij zelfs een regressie ten opzichte van de ideeën en principes van wijlen Ted Codd, grondlegger van het relationele model. Hij onderbouwt dit met zes argumenten. Het genoemde artikel 'Twijfels over logische correctheid' is een afgeleide van Gittens' publicatie op zijn website www.gittens.nl, zie [8].

Hugh Darwen en C.J. Date hebben Gittens' commentaar uiterst serieus genomen en krijgen van DB/M de gelegenheid hun standpunten en meningen ten aanzien van Gittens' opmerkingen diepgaand toe te lichten. Leerzame stof over de basisregels van het relationele database-model. Dit is de vierde bijdrage in een serie.

Verdere bespreking en uitwerking van de RM Voorschriften worden gegeven in Hoofdstuk 7 van TTM.

In het kort gaat RM Proscriptie 4 over het volgende. Per definitie bevatten tupels, en dus ook relaties, geen nullen (nullen zijn geen waarden!). SQL echter staat nullen toe in de tabellen – opnieuw het bewijs dat SQL-tabellen geen echte relaties zijn. In tegenstelling tot SQL verbiedt TTM het gebruik van nullen absoluut, categorisch en ondubbelzinnig (en is dus ook daarom n-waardige logica voor elke $n > 2$).

Er is een belangrijk punt dat eerst helder moet zijn voordat we onze afwijzing van nullen en driewaardige logica verder uitleggen en antwoord geven op Gittens' kritiek en suggesties betreffende dit onderwerp. Er is aangetoond (bijvoorbeeld in referentie [4] hoe het probleem dat door NULL wordt opgelost ook met een TTM-conform DBMS kan worden opgelost, zelfs zonder dat toevlucht hoeft te worden genomen tot 'speciale waarden'. Het is dus niet zo dat een probleem dat als opgelost wordt beschouwd door het gebruik van SQL's NULL, niet kan worden opgelost met D.

We vonden Gittens' kritiek op RM Proscriptie 4 moeilijk te volgen. Door het hele stuk [8] heen blijkt hij het helemaal eens te zijn met Codd – in het bijzonder met referentie [1] – en schrijft hij in paragraaf 6 dat Codd zelf in [1] het gebruik van nulls toestaat op een manier die veel weg heeft van SQL. Aan de andere kant lijkt Gittens het ook eens te zijn met onze afwijzing van SQL's driewaarden logica en met ons absolute verbod op het gebruik van NULL:

Date en Darwen lijken het concept van 'de onbekende' af te wijzen gebaseerd op problemen die men is tegengekomen in SQL. Maar ondanks dat velen het erover eens dat nulls geen waarden zijn, kan het toch best mogelijk zijn een taal met consistente semantiek te ontwerpen die om kan gaan met het concept van 'de onbekende' in de database, terwijl deze taal op het zelfde moment geen notie heeft van een instance van een null-waarde. Zo'n taal zou Codd's concept van een attribuut met onbekende waarde ondersteunen zonder de driewaarden logica te onderschrijven.

Na wat verdere uitwerking van dit idee, sluit hij af met de volgende, in cursieve letters gestelde vraag:

Geven Date en Darwen nu feitelijk een goede motivering voor de afwijzing van nulls?

De duidelijke tegenspraak in de geciteerde tekst maakt het moeilijk voor ons om te reageren en als we dat zouden doen, waaróp dan precies. Het is duidelijk dat Gittens onze rechtvaardiging voor de afwijzing van iets wil zien. In de laatste vraag noemt hij dat 'iets' expliciet 'nulls', maar in de eerdere tekst lijkt hij te vragen waarom we een andere benadering afwijzen, die helemaal geen nulls gebruikt en geen gebruik maakt van driewaarden logica. Dat zo'n benadering "Codd's concept van een attribuut met onbekende waarde zou ondersteunen" klinkt ons niet logisch in de oren, omdat Codd's concept zeker WEL gebruik maakte van nulls en zeker WEL gebruik maakte van driewaarden logica.

Hoe het ook zij, we zijn het er mee eens dat TTM geen onderbouwing voor onze afwijzing van SQL's NULL en SQL's driewaarden logica bevat. Dat komt deels doordat TTM een 'proposal' is die beoordeeld moet worden op de eigen merites (het boek [7] telt op dit moment 500 pagina's), en deels omdat onze gedetailleerde onderbouwingen voor het afwijzen van verschillende aspecten van SQL in het algemeen, en NULL in het bijzonder, in de afgelopen jaren veelvuldig op andere plaatsen verschenen zijn.

Ook beamen we dat TTM geen grond bevat om Gittens' idee af te wijzen, maar we vinden het onredelijk om zo'n motivering te verwachten. Waarom zouden we ook maar ergens een motivering geven voor de afwijzing van iets dat, voor zover we weten, nog nooit is uitgesproken?²

In het voorliggende artikel beantwoorden we eerst Gittens' vraag met een samenvatting van de onderbouwingen die we eerder publiceerden over de verwerping van SQL's nulls (en die van Codd uit 1979) en driewaarden logica. Vervolgens geven we commentaar op de door Gittens' voorgestelde benadering, waarbij het gebruik van zaken zoals NULL wordt vermeden en wordt vastgehouden aan tweewaarden logica.

Wat is er mis met driewaarden logica en NULL?

Als antwoord op Gittens' vraag "Geven Date en Darwen nu feitelijk een goede motivering voor de afwijzing van nulls?" geven we 18 hoofdstukken in de boeken zoals vermeld in referentie [5] en een ander in referentie [6].

We hebben nog een opmerking over wat Gittens "afwijzing van nulls" noemt in deze vraag. Dat is wat misleidend. Wij geven er de voorkeur aan te zeggen dat TTM een schema voorschijft gebaseerd op de wiskundige theorie van *n*-aire relaties, net als Codd deed in 1969. Die wiskundige theorie is op zijn beurt gebaseerd op de gevestigde en wijdverspreide First Order Predicate Calculus, die ook de 'propositional calculus' bevat die aantoonbaar al 2300 jaar onder ons is en waarvan de logica tweewaardig is. Het belangrijkste doel van TTM was om duidelijk te maken wat ervoor nodig is om een DBMS te maken

dat echt relationeel is. We vonden het niet passen om opnieuw uit te leggen waarom relationele databases verkozen moeten worden boven databases gebaseerd op andere ideeën, zoals hiërarchieën, netwerken, variabelen van elk denkbare soort (zoals in de objectgeoriënteerde benadering), of het soort structuren waaraan Gittens denkt. We vinden dan ook dat Gittens' vraag als volgt geparafraseerd kan worden: "Geven Date en Darwen nu feitelijk een goede motivering voor de afwijzing van andere structuren dan relaties?" En we zouden op dié vraag antwoord kunnen geven met de wedervraag "Geeft Gittens nu feitelijk een goede motivering voor de afwijzing van de pure relationele benadering?" Enfin, we geven maar weer eens onze mening over driewaarden logica. Onze mening over vier-en-hoger-waardige logica is soortgelijk, maar dan sterker (hoe meer waarden van de waarheid, hoe erger het wordt, zo zeggen wij).

De 'propositional calculus' die het hart vormt van ons fundament bezit verschillende belangrijke eigenschappen, zoals regel 1: Er zijn maar twee waarden van de waarheid, WAAR en NIET WAAR, aangevuld met nog acht daaraan verbonden regels (die u kunt lezen in de originele Engelse tekst op onze website, red.).

In [1], Codd's proposal uit 1979, komt geen van deze voorwaarden voor. In plaats daarvan gaat hij uit van: Er zijn drie waarden van de waarheid, WAAR, NIET WAAR en ONBEKEND. Ook alle daaraan verbonden regels wijzigen daardoor ingrijpend ten opzichte van de tweewaarden logica. (Zie onze website, red.). Codd verwierp zijn eigen voorstel uit 1979 ten faveure van een ander, gebaseerd op vier waarheidswaarden. Nadat hij het idee voor het eerst opperde in [2], publiceerde hij twee verschillende verzamelingen van vierwaarden logica waarheidstabellen, de eerste in [3] en de tweede, die de eerste ongedaan maakte, in een herdruk van dat boek in juli 1991. Meer details worden gegeven in [6], waar Date nog een later door Codd herschreven tekst beschrijft, die niet gevonden kan worden in enig algemeen verkrijgbare publicatie³. Het lijkt erop dat mensen die de drie- of vierwaarden logica aanhangen erg voorzichtig moeten zijn met het citeren van Codd ter ondersteuning van hun standpunt. In elk geval is ons commentaar op Codd's driewaarden logica met nóg meer kracht, *mutatis mutandis*, van toepassing op Codd's vierwaarden logica. Wat NULL betreft, de ondersteuning van dat concept impliceert per definitie ondersteuning van de driewaarden logica, dus verwerpen daarvan impliceert ook de verwerping van NULL. Vierwaarden logica komt ter sprake als twee verschillende (maar even schadelijke) varianten van NULL worden ondersteund.

Gittens stelt nog een tweede vraag waarop we willen reageren: *Bewijst The Third Manifesto eigenlijk wel dat er grondig onderzoek is gedaan naar de kwestie van de nulls?*

Het antwoord luidt 'nee'. TTM is een propositie, een voorstel, en geen commentaar op andere proposities. Wij vinden niet dat het tot de verantwoordelijkheid van TTM behoort om argumenten

SUPPLIER ID	NAME	CITY (mogelijk null)
S1	John	
S2	Jane	Paris
S3	Judy	

Afbeelding 1: Gittens' voorstelling van een relatie R in Codd-stijl.

aan te voeren tegen alles waar het boek niet over gaat. Het moet worden beoordeeld op zijn eigen kwaliteiten.

Gittens' suggestie

We kijken nog eens naar de suggestie van Gittens:

... kan het toch best mogelijk zijn een taal met consistente semantiek te ontwerpen die om kan gaan met het concept van 'de onbekende' in de database, terwijl deze taal op hetzelfde moment geen notie heeft van een instance of een null-waarde. Zo'n taal zou Codd's concept van een attribuut met onbekende waarde ondersteunen zonder de driewaarden logica te onderschrijven.

We zullen laten zien waarom wij denken dat dat geen goede suggestie is. Laten we beginnen met "geen notie heeft van een instance of een null-waarde." We zijn het er allemaal over eens dat NULL geen waarde is (omdat het niet volledig gelijk is aan zichzelf), nemen we aan dat deze zinsnede synoniem is aan "geen notie van een voorkomende NULL" (we maken indien dat nodig is onderscheid tussen het ergens voorkomen van een waarde en de waarde zelf). Maar kan iets betekenisvol bestaan zonder overal voor te kunnen komen? Wij denken van niet. We nemen dus aan dat Gittens eigenlijk bedoelt "geen notie van NULL." Hij illustreert zijn idee zoals in afbeelding 1. Vermoedelijk verwijst 'Codd-stijl' naar het voorstel uit 1979 in [1]. In dat geval representeren de lege cellen in deze tabel zeker een verschijningsvorm van NULL, en de kolom met de naam "CITY (mogelijk null)" staft deze interpretatie. We waren al enigszins in de war gebracht, maar dat wordt alleen maar erger als Gittens (die ons elders in [8] bekritiseert wegens een zeker gebrek aan rigoreusheid dat hij waarneemt in TTM) verder gaat met:

Zolang een taal kan garanderen dat de operators in die taal niet toestaan dat er relatiewaarden worden afgeleid uit onze voorbeeldrelatie R die verwijzen naar een onbekende, is alles in orde. Hetgeen in ons voorbeeld wil zeggen dat het statement

```
select * from R
```

een error zou opleveren.

Als er al iets zou bestaan als een "relatie" die "verwijst naar een onbekende", en als R zo'n "relatie" zou zijn, dan moet er op z'n minst één operator zijn die kan worden gebruikt om zo'n 'relatie' aan te duiden. Want R moet het resultaat zijn van het evalueren van een expressie die dat aanduidt!

We verdenken Gittens er sterk van dat hij eigenlijk bedoelt dat R

de huidige waarde is van een relatievariabele, zeg S. In dat geval moet R op de een of andere manier aan S zijn toegewezen. De expressie die in dat geval naar R verwijst is in het algemeen een aanroep door een bepaalde operator. Dat is zelfs het geval als de opdracht aan S heeft plaatsgevonden door het gebruik van de bekende verkortingen INSERT, UPDATE of DELETE. Als voorbeeld: `DELETE S WHERE City = 'Paris'` is kort voor de opdracht: `S := S MINUS (S WHERE City = 'Paris')`

De expressie aan de rechterkant van deze opdracht is de aanroep van een operator, MINUS, die normaliter gedefinieerd is om uit twee relaties één relatie te verkrijgen. Gittens wil onderscheid maken tussen 'relaties' die op de een of andere manier "verwijzen naar een onbekende" (laten we dat RVO's noemen) en relaties die dat NIET doen (en die we gewoon 'relatie' blijven noemen). De bovenstaande aanroep van MINUS zal zeker iets uit R afleiden, waarbij R de huidige waarde van de variabele S is waarnaar twee keer wordt verwezen in de aanroep. Bovendien is een van de taken de aanroep van de operator WHERE, die ook iets uit R zal afleiden. En wat te denken van de eerste taak van MINUS, die alleen maar 'S' is? Dat is een variabele verwijzing, een expressie die als die wordt uitgevoerd de huidige waarde van S zal opleveren. Logischerwijs is dat ook een aanroep van een operator. Als het die operator niet is toegestaan om een RVO als resultaat te leveren, en de huidige waarde van S is een RVO en geen relatie, dan is de expressie 'S' volgens Gittens' idee verboden – iets wat eveneens geldt voor de expressies `S WHERE City = 'Paris'` en `S MINUS (S WHERE City = 'Paris')`.

Deze zojuist gedane constatering is volkomen consistent met Gittens' suggestie dat de SQL expressie `select * from R` een error tot gevolg heeft. (Feitelijk is 'Select * from R' SQL's idiosyncratische lange manier voor wat in de meeste talen geschreven zou worden als alleen 'R'.)

We zullen Gittens verder ontzien en de lezer verdere pijnlijke ontkrachting van zijn suggestie besparen. Meer opbouwend

Dit is een bewerkte en vertaalde versie van de originele Engelse tekst, die u kunt vinden op onze website www.dbm.nl onder 'specials', 'extra materiaal'. In geval van discussies is de originele Engelse tekst doorslaggevend.

doen we de suggestie dat als een taal het onderscheid tussen RVO's en relaties moet kunnen maken zoals Gittens in gedachten heeft, dat die taal dan minimaal al het volgende nodig heeft:

- operators die als ze worden aangeroepen relaties leveren;
- operators die als ze worden aangeroepen RVO's leveren;
- operators die taken uitvoeren op relaties;
- operators die taken uitvoeren op RVO's;
- ondersteuning van relatie-variabelen en operators om die variabelen te updaten (inclusief opdrachten);
- ondersteuning van RVO-variabelen en operators om die variabelen te updaten (inclusief opdrachten);
- vergelijkingsoperators voor relaties;
- vergelijkingsoperators voor RVO's.

Met The Third Manifesto en de taal Tutorial D om de propositie te illustreren, hebben we de database-gemeenschap iets gegeven dat kan worden geëvalueerd, bekritiseerd, geaccepteerd of verworpen. Als Gittens ons op dezelfde manier had voorzien van een definitie of een model gebaseerd op zijn suggestie en een voorbeeld van een taal die zijn ideeën omarmd, zouden we op diezelfde manier in de gelegenheid zijn geweest dat te evalueren enzovoort.

Want eigenlijk denken we dat het een erg goed idee is om zo'n model te definiëren, al is het alleen maar om iets te evalueren te hebben. Tot dat moment hebben we over deze kwestie niets meer zeggen.

Noten

1. *The Third Manifesto* is de titel van Hoofdstuk 4 van *The Third Manifesto* en de afkorting verwijst eigenlijk alleen naar dat hoofdstuk.
2. Feitenlijk kan het paper van wijlen Adrian Larner [9] worden gezien als zo'n voorschrift, maar het paper kreeg niet veel aandacht en blijft volgens de huidige auteurs niet overeind bij zorgvuldige analyse.
3. Zie hoofdstuk 18, 'Waarom drie- en vierwaarden logica niet werkt'.

Literatuur

1. E.F. Codd: *Extending the Database Relational Model to Capture More Meaning*, ACM TODS 4, No. 4, December 1979.
2. E.F. Codd: *Missing Information (Applicable and Inapplicable) in Relational Databases*, ACM SIGMOD Record 15, No. 4, December 1979.
3. E.F. Codd: *The Relational Model for Database Management, Version 2*, Reading, Mass.: Addison-Wesley 1990.
4. Hugh Darwen: *How to Handle Missing Information without Using NULL*, Presentation slides available at www.thethirdmanifesto.com.
5. C.J. Date: *Relational Database: Selected Writings*, Reading, Mass.: Addison-Wesley (1985 – 1998).
6. C.J. Date: *Date on Database, Writings 2000-2006*, Berkeley, Ca.: Apress, 2006.
7. C.J. Date and Hugh Darwen: *Databases, Types, and the Relational Model: The Third Manifesto (3rd edition)*. Boston, Mass.: Addison-Wesley (2006).
8. Maurice Gittens: *The Third Manifesto Revisited* www.gittens.nl/TheTTMRevisited.pdf.
9. Adrian Larner. *A New Model of Data* at www.marmaladesky.co.uk/database.

Hugh Darwen en **Chris Date** zijn auteurs van de derde editie van *The Third Manifesto*.



Do you have a Point?

Do you make your Point?

Do you get the Point?

Do you want to prove your Point?

Join FourPoints !!