

# On logical mistakes in The Third Manifesto

Maurice Gittens <maurice at gittens dot nl>

September 12, 2006

## Abstract

According to the authors, Hugh Darwen and C.J. Date, of the third edition of the Third Manifesto entitled “Databases, types and the relational model; The Third Manifesto” the maxim: *All logical difference are big differences* and its corollary *All logical mistakes are big mistakes* has been central to their work on this book. Respecting the standard set by this maxim and its corollary, this paper will proceed to identify a number of issues with the logical consistency of the dissertation presented in The Third Manifesto. The conclusion is drawn that the Date and Darwen are regressing relative to Codd’ RM/T work.

The copyright of this document belongs to its author. Making complete and unmodified copies of this document is allowed.

**Status: draft**

## Revision history

- Sept 12 2006; Minor editing
- Sept 8 2006; Added section on nulls and relation valued attributes
- Sept 7 2006; Add section about “Treating operators as relations”
- Sept 6 2006; A first draft of this document

## Contents

<b>1 No more Great Blunders</b>	<b>2</b>
<b>2 Treating operators as relations without proper substantiation</b>	<b>2</b>
<b>3 No adherence to the principle of semantic compositionality</b>	<b>3</b>
<b>4 No semantic integrity in the presence of relational assignment</b>	<b>3</b>
<b>5 Undermining issues with relation valued attributes</b>	<b>5</b>

<b>6</b>	<b>No sound substantiation for the rejection of null</b>	<b>6</b>
<b>7</b>	<b>Conclusions</b>	<b>7</b>

## **1 No more Great Blunders**

Relative to the second edition of The Third Manifesto, the third edition has changed, on many points. On the positive side Date and Darwen have reduced the polemics and are generally more precise in their definitions. So, this edition gives a better impression of their position on databases, types and the relational model as opposed to the previous version which put to much effort into describing what they are opposed to. There is no longer much ado about the so-called *great blunders*, so prominently present in the second edition of the The Third Manifesto.

However, there remain issues with the logical consistency of the dissertation presented in the third edition of The Third Manifesto. In my opinion, the standard set by the maxim: *All logical difference are big differences* and its corollary *All logical mistakes are big mistakes* that Date and Darwen present as a guiding principle in their work, is not met by The Third Manifesto. Important issues which have influenced my position are:

- Treating operators a relation without proper substantiation
- No adherence to Frege's principle of semantic compositionality
- No semantic integrity in the presence relational assignment
- Undermining issues with relation valued attributes
- No sound substantiation for the rejection of null

An elaboration on these issues will be presented in the following.

## **2 Treating operators as relations without proper substantiation**

In appendix A of The Third Manifesto, Date and Darwen, repeatedly make the unqualified statement that operators can be treated as relations. This idea, which they claim as their own<sup>1</sup>, is used as ground for the dispensing with a few operators from Codd' algebra as can be read in the following quote taken from Appendix A.

We now claim that, given the fact that operators can be treated as relations, and given also the availability of the A operators AND, REMOVE, and RENAME (the latter two still to be discussed), it is indeed the case that we can dispense with restrict, EXTEND, and SUMMARIZE. We will justify this claim in the next section but one.

---

<sup>1</sup>See page 366

First, it seems fair to assume that many people familiar with truth tables as they are known in the context of boolean algebra, have for many years appreciated the fact that *commutative* operators can be treated as relations. So, Date and Darwen seem to have discovered hot water here. Second and more significantly, while making claims about operators in general, Date and Darwen attempted to motivate their dispensing specific operators from Codd' algebra based on a simple example involving a *commutative* operator PLUS. Rigour would require them to show that pertinent operators from Codd' algebra are also commutative. Alternatively, rigour would require an elaboration on noncommutative operators such SUBTRACT as well. Attempting to view noncommutative operators as relations one soon encounters the problem that it is necessary not only to identify the operands of these operators but it is also necessary to designate their respective roles. Information about the roles of operands is simply lacking in the relational representation of these operators and Date and Darwen would need to illustrate how this information could be catered for without violating their proscriptions and prescriptions.

### **3 No adherence to the principle of semantic compositionality**

The type of a relvar is determined by the header of the relvar. The candidate keys associated with a relvar are *not* part of their type according to The Third Manifesto.<sup>2</sup>This choice by Date and Darwen, represents a serious logical error because it causes variables and values of the *same* type to not be interchangeable. This is evident when one considers a relation value  $C$  of type  $T$  may not be assignable to a relational variable  $V$  of type  $T$ . More specifically, the assignment of  $C$  to  $V$  is not allowed when there are candidate key constraints defined on  $V$  to which  $C$  is not in adherence. Put another way, even though  $V$  and  $C$  share the same type, the assignment  $V = C$  may or may not work depending on whether or not  $C$  is in adherence with all candidate key constraints defined for  $V$ . Similarly, relation constants, relation nested values and non nested relation values, cannot play the role of *parent* in foreign-key relations using facilities provided by the The Third Manifesto. This is obviously true because, according to The Third Manifesto, relation constants and relation values have no<sup>3</sup> associated candidate keys and foreign-keys are defined in terms of candidate keys of the parent relation variable.

### **4 No semantic integrity in the presence of relational assignment**

Breaches of semantic integrity invalidates trust in data stored, thus seriously compromises value. Consider the relvar predicate for a relation Author with one candidate key (based on surname) and two tuples.

---

<sup>2</sup>See prescription 13 and 15 on page 85, point 13.

<sup>3</sup>The TTM does assign implicit candidate keys to virtual relvars. But not to relation values and relation constants.

SURNAME	firstname
Date	Chris
Darwen	Hugh

The *relvar predicate* (see page 29) for this relation might read:

Author with surname "Date" has first name "Chris".  
 Author with surname "Darwen" has first name "Hugh".

Consider multiple assignment where the key values for both tuples are swapped.

```
update Author where surname = "Date" { surname = "Darwen" },
update Author where surname = "Darwen" { surname = "Date" };
```

Since integrity checking is postponed until after all assignment have been performed<sup>4</sup>, this assignment statement is allowed by the The Third Manifesto. The relation value of V now becomes.

SURNAME	firstname
Darwen	Chris
Date	Hugh

Which reads:

Author with surname "Darwen" has first name "Chris".  
 Author with surname "Date" has first name "Hugh".

Given the relation values of the *Author* relvar *before* and *after* this single assignment and the knowledge that only one assignment has taken place, consider an forensic application which needs to find out what was *changed* by this assignment statement. Not appreciating this fundamental breach of integrity facilitated by The Third Manifesto will likely lead to seriously erroneous conclusions like:

The first name of the Author with surname "Date" was changed to "Hugh"

and

The first name of the Author with surname "Darwen" was changed to "Chris".

What, could in general be done to avoid more realistic<sup>5</sup> variants of this example to cause compromise of the semantic integrity of relational databases adhering to the The Third Manifesto? In essence the root of the problem is the fact that the definition of identity and assignment in the The Third Manifesto lacks facilities for specifying immutability. So, while The Third Manifesto transition constraints would allow one to detect change of relvars, there is no way to specify immutability of key values in assignments. So Date and Darwen cannot have it both ways. So they need to choose:

<sup>4</sup>See page 179.

<sup>5</sup>If you are still having doubts, please consider a variant of the above based around bank account numbers. Imagine a bank with a Date and Darwen styled database. What would happen in a financial system handling hundreds of transaction per second, if it were to loose track of account numbers, by means similar to those presented above?

1. or, we need to identify tuples changed by assignment statements.
  - (a) Codd' immutable, system assigned, surrogates serve this purpose, or
  - (b) facilities need to be provided to specify which keys are immutable in an assignment
2. or, we need to accept that relvar predicates are not a suitable vehicle for expressing propositions represented by relation variables as explained on page 29.

The first option entails the recognition of tuple identity and the second option entails letting go of the notion that relvars correspond with a set of propositions. However Date and Darwen choose to address this issue, fact remains that it represents a serious regression relative to Codd' RM/T because in the presence of a *single* assignment statement, the semantic integrity of a complete database can be compromised.

**Consequences** Violation of the semantic integrity, as facilitated by The Third Manifesto' lack of support for tuple identity, is likely to be among the most severe forms of compromise for an information system, because it compromises trust. Codd' immutable surrogates<sup>6</sup> allows relation variables to be understood as sets of tuple variables, thus providing the ability to trace change to individual tuple variables. Consequently, Codd style databases facilitate for example, accountability, tracibility and forensics in a way not possible with Date and Darwen style databases. From the perspective of information security, I dear to say that the position on relational assignment and tuple identity as held by Date and Darwen is naive. Codd presented a solution in 1979! He introduced the concept of the immutable tuple identifier, upon which semantic integrity can be maintained in a conceptually pure manner from the perspectives of logic and information security.

## 5 Undermining issues with relation valued attributes

Contrary to Codd, Date and Darwen allow relation value attributes. Please consider the following questions.

- What problem is solved by support for these attributes that could not be solved otherwise?
- What propositions can be represented by relvars including relation values attributes, that cannot be represented, by other relational means?
- What positive traits of alternative solutions are not available to solutions based on relation valued attributes?

---

<sup>6</sup>For the purposes of semantic integrity and information security, in the presence of change, there is at least one requirement in addition to immutability.

These questions are specifically pertinent in the light of the fact that these relation valued attributes are *relation values* and as such they lack associated candidate keys. In addition relation valued attributes cannot play any role in foreign-key constraints given the facilities provided by The Third Manifesto. Which is to say that alternatives to relation valued attributes have more facilities to accurately constrain databases to adhere to requirements of business and other applications.<sup>7</sup>

So, the question remains, why should the relational model be complicated for dubious gain relative to Codd's alternative? Appendix B of The Third Manifesto, represents an elaboration of sorts on this topic. This chapter digresses much but provides little that is of substance. I quote<sup>8</sup>:

What then is the criterion for making something a type and not a relvar?  
In our opinion this question is still somewhat open.

Put another way Date and Darwen do not seem to know, in any definite sense, what the advantages of relation valued attributes are. Relative to alternative relation solutions, the disadvantages however, are clear:

the employment of relation valued attributes as introduced by The Third Manifesto provide less opportunity for the expression of candidate key and foreign key constraints in databases, relative to alternatives not involving relation valued attributes.

Is it not up to Date and Darwen to provide proper arguments for adding relation valued attributes to the relational model? Until clear advantages of relation valued attributes can be illustrated, relative to alternatives, support for such attributes seem a needless complication.

## 6 No sound substantiation for the rejection of null

Nulls in databases are a source of much discussion. Codd, together with many practitioners allow nulls to be used where pragmatically deemed appropriate. Date and Darwen disagree. This edition of the Third Manifesto provides little substantiation for their position. This is a pity because it would have been nice to have a solid presentation of their arguments which amounts to more than ranting against SQL. For example, since many agree that nulls are not values, it just might be possible that a language with consistent semantics can be designed that accommodates the concept of nulls in databases while the same language has no notion of an instance of a null value. Such a language would accommodate both Codd's concept of an *unknown* value and Date and Darwen's rejection of instances of *null* values. Consider for example the following proposition corresponding to an example relvar predicate taken from The Third Manifesto<sup>9</sup>.

---

<sup>7</sup>For example, the GROUP statement does not provide facilities for specifying candidate key constraints and foreign key constraints

<sup>8</sup>See page 379

<sup>9</sup>See page 29

Supplier *S1* is under contract, is named *Smith*, has status *20*, and is located in city *London*.

Under the assumption that the city attribute of the corresponding relvar accommodates nulls and that the city corresponding with supplier *S1* is unknown, here are at least two options for the corresponding proposition:

- Supplier *S1* is under contract, is named *Smith*, has status *20*.
- Supplier *S1* is under contract, is named *Smith*, has status *20*, and is located in an *unknown* city.

What would a language exploiting the first option look like? Cannot conceptually pure and consistent, language semantics be devised based on this first proposition? The Third Manifesto has provided no evidence that such alternatives have been pursued to a degree that informed choices were made on the matter. Or would Date and Darwen have us assume that Codd did not have a proper understanding of things relational?

## 7 Conclusions

The relational model as presented by Date and Darwen in the third edition of The Third Manifesto represents a regression relative to Codd RM/T on the following grounds:

- “Treating operators as relations” approach does not seem to be well researched.
- The Third Manifesto does not facilitate the semantic integrity of databases in the presence of a single assignment. This is likely to facilitate many forms of information security breaches based on identity theft.
- The Third Manifesto is inconsistent with regards to assignment of value and variables of the same type. More specifically, because the type of relation variables does not include their associated candidate keys, variables and values of one and the same type may not be assignable to each other. Consequently, The Third Manifesto is not in adherence with the Frege’ principle of semantic composition.
- Since, no clear *added value* is to be associated with relation valued attributes and the existence of disadvantages relative to alternatives, support for relation values attributes does not seem to be warranted.
- Nulls rejected without sound substantiation.

## References

- [1] L.T.F Gamut [1991] Logic, Language and Meaning Volume 2, University of Chicago Press.
- [2] E.F. Codd, [1979] “Extending the Database Relational Model to Capture more Meaning”

- [3] C.J. Date, Hugh Darwen [2007] Databases, Types and the Relational model, Addison-Wesley Publishing Company.
- [4] C.J. Date, Hugh Darwen [2000] Foundation for Future Database Systems, Addison-Wesley Publishing Company.
- [5] Maurice Gittens [2003], "A critical reading of the third manifesto", <http://www.gittens.nl>