Appeared in SIGMOD RECORD, March 1993, pp. 57-58.

Response to "Remarks on two new theorems of Date and Fagin"

C. J. Date
Independent Consultant

Ronald Fagin
IBM Almaden Research Center

In [DF92], we present simple conditions, which we now describe, for guaranteeing higher normal forms for relational databases. A key is *simple* if it consists of a single attribute. We show in [DF92] that if a relation schema is in third normal form (3NF) and every key is simple, then it is in projection-join normal form (sometimes called fifth normal form), the ultimate normal form with respect to projections and joins. We also show in [DF92] that if a relation schema is in Boyce-Codd normal form (BCNF) and *some* key is simple, then it is in fourth normal form (4NF). These results give the database designer simple sufficient conditions, defined in terms of functional dependencies alone, that guarantee that the schema being designed is automatically in higher normal forms.

In [Bu93], Buff gives a nice generalization of the second of these results. He defines a set C of attributes to be a *cut* of a relation schema if every key of the schema intersects both C and its complement. He proves that if a relation schema is in BCNF and has no cut, then it is in 4NF. Buff's Theorem is an immediate consequence of the lemma (Lemma 5.1) that we used in [DF92] to prove our second result. Therefore, we should have discovered it (but we didn't!)

In addition to giving this result, Buff makes comments on the usefulness of such theorems. He claims that such theorems are not especially useful to the practitioner, for the reason that these theorems are not "extendible". What he means by this is that the properties we consider (such as "the schema is in 3NF and every key is simple") may no longer be true if a new constraint is added to the schema. We do not accept this criticism as valid. Adding a new constraint to a schema is a drastic change. It is clear that a schema that is "good" (say, in some higher normal form) can become bad if a new constraint (such as a new functional dependency) is added to the schema.

Buff seems to imply that for a property to be robust, it must continue to hold whenever a constraint is added. Why not also say that for a property to be robust, it must continue to hold whenever a constraint is dropped? In fact, it is arguably more likely in practice

for a constraint to be dropped than for a constraint to be added. For example, we might suddenly decide to allow an employee to have more than one telephone number, which amounts to dropping the functional dependency that says that each employee has only one telephone number. On the other hand, if a constraint were *added*, this might invalidate current data. Let us define a property to be *robust* if it continues to hold whenever a constraint is either added or dropped. Let us say also say that a property P is *interesting* if there are two schemas $\mathcal{S}$ and $\mathcal{S}'$ with the same set of attributes, such that $\mathcal{S}$ satisfies property P, but $\mathcal{S}'$ does not satisfy property P. As we now show, there does not exist any robust property that is interesting! This is because if $\mathcal{S}$ and $\mathcal{S}'$ are arbitrary schemas with the same set of attributes, then we can obtain $\mathcal{S}'$ from $\mathcal{S}$ by first dropping constraints from $\mathcal{S}$ one at a time until the set of constraints is empty, and then adding constraints one at a time until we obtain the constraints in $\mathcal{S}'$. (We are taking advantage of the usual assumption that each schema has only a finite number of constraints.) So if $\mathcal{S}$ satisfies a robust property P, then so does $\mathcal{S}'$. This shows that this notion of robustness is not reasonable. Similarly, by our comments above, Buff's closely related requirement of extendibility is not reasonable.

We feel that Buff misses the point of our paper. As we say in [DF92]:

> These results give conditions that are easy for the practitioner to understand and that are sufficient to guarantee the higher normal forms. Thereby, they provide a practical database design guideline, that may make the database designer's job a little easier. These results are also useful for the database instructor, who can give the class practical situations in which projection-join normal form can be achieved, without requiring knowledge of multivalued dependencies or join dependencies.

We have been told that in fact, students are enthusiastic about the results in our paper, because they are happy to get simple-to-understand sufficient conditions for "goodness" of a relation schema.

References

[Bu93] H. W. Buff, Remarks on two new theorems of Date and Fagin. *SIGMOD Record*, March 1993.

[DF92] C. J. Date and R. Fagin, Simple conditions for guaranteeing higher normal forms for relational databases. *ACM Trans. on Database Systems* 17,3, Sept. 1992, pp. 465–476.