XSP: Extended Set Processing

Mathematically Managing Data Representations

D L Childs
<iis@umich.edu>

1. INTRODUCTION

Computer applications process *data representations*. The form, content, and availability of data representations limit or assist the performance of applications.

This paper presents a high level overview of why a mathematical model of data representations is necessary, and how an extended set processing model accomplishes the task.

An ideal data processing environment would be able to supply each application with just the required relevant data, in just the required format, and at just the required time.

Since no single data representation is likely to be ideal for all phases of the data processing process, each phase must be considered separately along with the transformations between data representations.

2. DATA REPRESENTATION

A typical data processing system requires four unique and independent representations of data:
1) Data representation for Presentation.
2) Data representation for Processing.
3) Data representation for Propagation.
4) Data representation for Preservation.

The key problem is in choosing a model that preserves the data *content* within and between respective data representations.

3. MODELING ISSUES

The traditional approach has been to pick a model with a single data representation, such as a *Table* and use it for all four data representation needs.

Though this approach has had a serviceable success for over thirty years, the need for accessing large volumes of diverse, distributed data representations has challenged the performance capabilities of traditional systems.

3.1 Set Theory

It is generally acknowledged that a formal foundation provides benefits for managing a complex technology. It is also generally acknowledged that mathematics provides a most rigorous formal foundation. Since classical set theory, CST, is generally attributed to be a foundation for all mathematics, it might seem appropriate to choose CST as a foundation for modeling data representations. It is, in fact, a most inappropriate choice.

3.2 Content vs. Structure

Though the success of the Relational Data Model, RDM, would seem to contradict the choice of CST as being inappropriate, the RDM was (and still is) successful in spite of its use of CST, not because of it.

The advantage the RDM took of CST was the use of *set operations*, not the use of *sets*. Sets provide precise definitions for *content*, but no recognition of *structure*. Data representations require precise specifications of both content and structure.

3.3 Adding Structure to Sets

Extended set theory, XST[1], was developed to add a structural component to accompany the content component of *set membership*. For example:
1) In CST 'x is a member of A' is expressed as $\{x\}$.
2) In XST 'x is a y-member of A' is expressed as $\{x^y\}$.

Since XST places virtually no restriction of the values for 'y', y's could be integers, memory locations, file addresses, domain names, or even other sets.

Thus XST provides an excellent foundation for modeling the four different data representation types.

4. SET OPERATIONS

The next modeling issue concerns operations that can transform one data representation into another while preserving data content.

Since all data representations have a mathematical identity under the axioms of XST, any and all data processing can be expressed with extended set operations.[2]

An additional benefit of using extended (or structured) sets is the ability combine multiple executable processes into a single executable process, a *composition*[3] of processes.

5. CONCLUSION

Extended set processing systems have been in commercial use since 1972.[4] The technology is proven and available to anyone willing to tolerate the arcane notation.

6. REFERENCES

[1] Blass, Childs: *Axioms & Models for an Extended Set Theory*, http://www.math.lsa.umich.edu/%7Eablass/XST_Axioms.pdf
[2] Childs: *Why Not Sets?* http://xsp.xegesis.org/WNSETS.pdf
[3] Childs: *Functions as Set Behavior* http://xsp.xegesis.org/Process.pdf
[4] MICRO RDBMS (1972-1998) http://en.wikipedia.org/wiki/Micro_DBMS
[5] **XSP Technology**: (webpage), http://xsp.xegesis.org