Extended Subsets

Ben Laurie
(ben@links.org)

November 8, 2009

1 Introduction

Marsh Ray has a website called extendedsubset.com. I asked him what an extended subset was, and he explained that it was what happened to languages like Ada or PL/I. They were so large that no-one managed to implement all of them, but each vendor added their own proprietary extensions. The result is an extended subset.

It occurred to me that we can define extended subsets as an analogue to sets.

2 Extended Sets

Before you have a subset, it's a good idea to have a set. So, an extended set S^* is a set of sets, $(S_1, ..., S_n)$. Define

$$S = \cap S_i \tag{1}$$

In the case of Ray's example, S would be Ada or PL/I and the S_i would be various vendor's extensions.

3 Primitive Extended Subsets

A primitive extended subset is a pair, $P = (P_s, P_{S_i})$, where P_s is a subset of S and P_{S_i} is a subset of some S_i.

4 Extended Subsets

Finally.

An extended subset is a set of primitive extended subsets, $T^* = (P_i)$ s.t. each P_i has the same component, T, in S, as all the others.

So, in Ray's example, an extended subset would be a collection of implementations of Ada. T would be their common ground - the subset of Ada they all implemented, and each implementation's T_{S_i} would be their extensions and the parts of Ada that were standard but not common.

That last is an interesting point: because each S_i is a superset of S, T_{S_i} can contain elements from S which aren't in T.

5 Intersection and Union

Sets are no fun without operations, so how do these work?

The intersection of two extended subsets is easy. If $T^* = (P_i)$ and $U^* = (Q_i)$, then

$$T^* \cap U^* = ((T \cap U, P_{i_{S_k}} \cap Q_{j_{S_k}})) \tag{2}$$

Note that it is expected that there may not be any such thing as $P_{i_{S_k}}$ or $Q_{j_{S_k}}$ for some values of i and j, in which case, those elements do not exist in the intersection.

and likewise the union,

$$T^* \cup U^* = (T \cup U, P_{i_{S_k}} \cup Q_{j_{S_k}}) \tag{3}$$

6   Outer Intersection?

It seems obvious there are extra operations possible on extended subsets. In particular, the one I would want to create when comparing two implementations of Ada. This would be:

$$T^* \cap^* U^* = ((T \cap U, P_{i_{S_i}})) + ((T \cap U, Q_{i_{S_j}})) \tag{4}$$

which we might call outer intersection. This would combine the common subset of Ada with each set of non-common Ada and proprietory extensions.

7   Research Directions

What if $S_i \cap S_j = S$ is required?

8   Conclusion

What Ray described is actually a primitive extended subset. So, Ray's website should be primitiveextendedsubset.com.

9   What's The Point?

Nothing, really. I assumed an extended subset was a real thing, so, as is my habit, I tried to figure out what it's structure might be. This is what I came up with. I thought it was a fun enough toy to write it down.