Evaluating Invariances in Document Layout Functions

Alexander Macdonald

David Brailsford

John Lumley

Variable Data Printing

- Digital offset presses allow for short runs
- Each document potentially unique
- Pre-rasterisation not always feasible:
 - Double sided A4 page @ 600ppi 24bpp =
 180 MB
 - 500 page catalogue = 90 GB
 - 1,000,000 unique catalogs = 90 PB

Optimising Variable Data Documents

- Potentially unique, probably quite similar
- Analysis of document to find invariant sections
- Invariant sections can be rasterised for re-use
- Could be used to generate efficient PPML

Progressive Document Evaluation

- A document is a tree of nested functions and content
- Unbound items exist where future content will be placed
- As more items become bound more invariant sections will be discovered

Layout Functions



Layout Functions 2



Layout Functions 3



Dealing with Unbound Items

linear flows exhibit associativity:

$$f(a, f(b, c)) == f(f(a, b), c)$$

 Associativity allows successive bound items to be evaluated:

$$f(u, b, b, u, u, b, b) == f(u, f(b, b), u, u, f(b, b))$$

== $f(u, b, u, u, b)$

 A bound item at the head of a flow can be moved outside:

$$f(b, u, b, u) == b, f'(u, b, u)$$

Implementation

 SVG with elements from the "layout" namespace

- Layout processor takes two inputs:
 - Document to be processed
 - Data to replace an unbound item

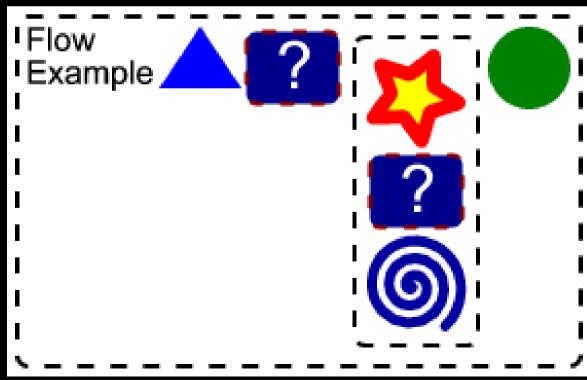
 After all unbound items have been bound the document is fully evaluated to renderable SVG

Example

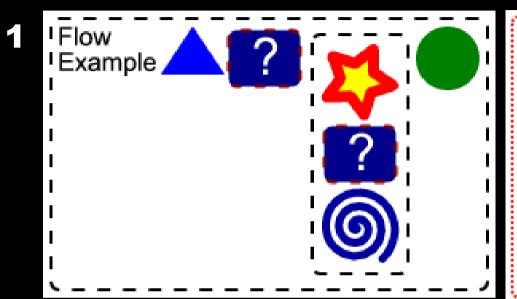
</lavout:yflow>

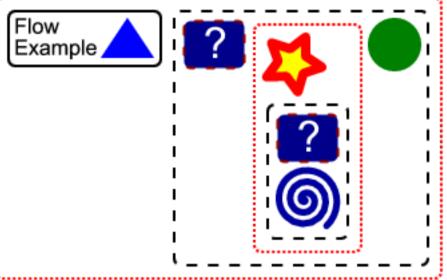
<svg:svg/>

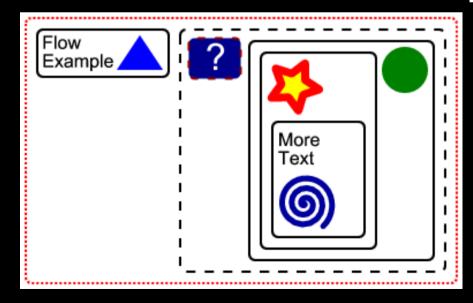
</la>

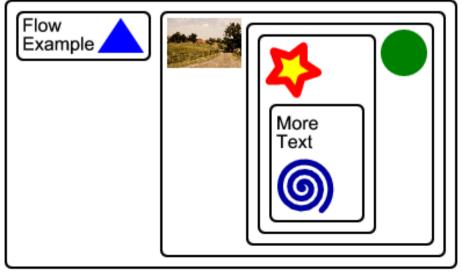


Example









3

Conclusion

 Progressive evaluation of documents is beneficial to VDP

 Future work will look at applying more program reasoning techniques from the FP community to documents

Questions?