

Data Representation Alternatives in Semantically Augmented Numerical Models

Mike Fagan Laurent Hascoet
Jean Utke

Takeaway

- What is Program Augmentation?
- What is the Result?

Program Augmentation

- Program transformation
- Opposite of 'Program Elision'
- Augment programs to compute additional mathematical artifacts
- Interdisciplinary: symbolic computation, SCAM
- Motivating Example: Automatic Differentiation (AD)

Program Augmentation

- Program transformation
- Opposite of 'Program Elision'
- Augment programs to compute additional mathematical artifacts
- Interdisciplinary: symbolic computation, SCAM
- Motivating Example: Automatic Differentiation (AD)

Automatic Differentiation

- Program computes $f(X)$, AD returns a program that computes both $f(X)$ and $f'(X)$
i.e. program is *augmented* with derivatives

- Sample:

$$y = z * w$$

$$dy = z * dw + dz * w$$

- Adifor (Fagan), Tapenade(Hascoet),
OpenAD(Utke)

Program Augmentation 2

- Augmentation requirements:
 - What to do with primitives
 - Arithmetic
 - Intrinsic functions
 - Sequencing method (= chain rule)

Other Augmentations

- Divided Differences
- Guaranteed bounds (=interval arithmetic)
- Stochastic moments
- Probability Distributions

What is the Result?

- During augmentation, where do you put the augments?
- There are 2 different 'styles':
 - By-name
For each variable v , $\text{aug}(v) = \text{new variable } y \text{ ó } dy$
 - By-address
next slide

By-Address Augment Representation

- For each variable, change the type to a 'structured' type
double ϕ struct _d {double v; double d[]}
- Now, augments are referenced via structure projection:
 $y.v = x.v * z.v$
 $y.d = x.v * z.d + z.v * x.d$
- (Note: hybridization possible for structured types)

What is the Problem?

- By-name has been thought to be problematic for C-like languages [e.g. C, C++, maybe F90]
- By-address has been thought to be problematic for languages without structured types [F77], and for languages with array slices* that do not commute over component reference [some F90, MATLAB]

`x[1:10]%v ???`

Our Result

- You can use by-name or by-address as you choose !
- By-name for C-like languages
[full or partial separation]
- By-address for F77 and MATLAB
[use array indexing instead of component reference]

Controversy

- Compose augmentations?
- Analyses improved?
 - Activity (like slicing)
 - Linearity
- Trace compression via program reversal?
[floating point is an issue]