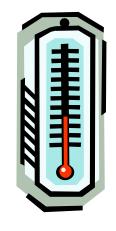




## Types for Units-of-Measure in F#



Andrew Kennedy Microsoft Research Cambridge





### NASA "Star Wars" experiment, 1983



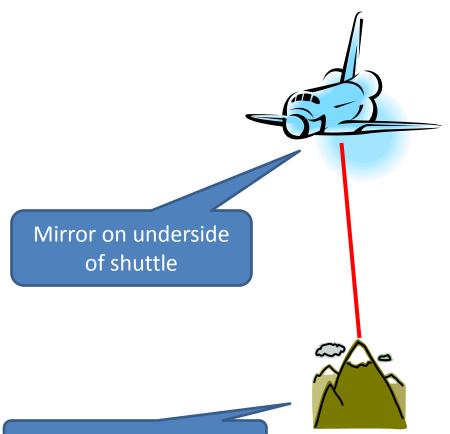
23<sup>rd</sup> March 1983. Ronald Reagan announces SDI (or "Star Wars"): ground-based and space-based systems to protect the US from attack by strategic nuclear ballistic missiles.





### 1985



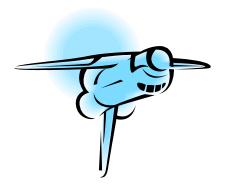


SDI experiment: The plan

Big mountain in Hawaii

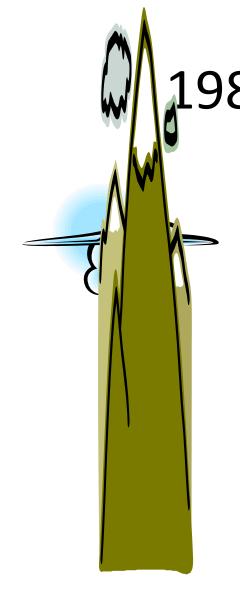
### 1985





SDI experiment: The reality







The reality

#### Attention All Units, Especially Miles and Feet!

Much to the surprise of Mission Control, the space shuttle Discovery flew upside-down over Maui on 19 June 1985 during an attempted test of a Star-Wars-type laser-beam missile defense experiment. The astronauts reported seeing the bright-blue low-power laser beam emanating from the top of Mona Kea, but the experiment failed because the shuttle's reflecting mirror was oriented upward! A statement issued by NASA said that the shuttle was to be repositioned so that the mirror was pointing (downward) at a spot 10,023 feet above sea level on Mona Kea; that number was supplied to the crew in units of feet, and was correctly fed into the onboard guidance system -- which unfortunately was expecting units in nautical miles, not feet. Thus the mirror wound up being pointed (upward) to a spot 10,023 nautical miles above sea level. The San Francisco Chronicle article noted that "the laser experiment was designed to see if a low-energy laser could be used to track a high-speed target about 200 miles above the earth. By its failure yesterday, NASA unwittingly proved what the Air Force already knew -- that the laser would work only on a 'cooperative target' -- and is not likely to be useful as a tracking device for enemy missiles." [This statement appeared in the S.F. Chronicle on 20 June, excerpted from the L.A. Times; the NY Times article on that date provided some controversy on the interpretation of the significance of the problem.] The experiment was then repeated successfully on 21 June (using nautical miles). The important point is not whether this experiment proves or disproves the viability of Star Wars, but rather that here is just one more example of an unanticipated problem in a human-computer interface that had not been detected prior to its first attempted actual use.

### NASA Mars Climate Orbiter, 1999



spacecraft operation, according to a review finding released Thursday.

The units mismatch prevented navigation information from transferring between the Mars Climate Orbiter spacecraft team in at Lockheed Martin in Denver and the flight team at NASA's Jet Propulsion Laboratory in Pasadena, California.

multimedia showcase

more services

#### E-MAIL:

Subscribe to one of our news e-mail lists. Enter your address:

#### Solution

- Check units at development time, by
  - Static analysis, or
  - Type checking

#### **Chapter 18** Validating the Unit Correctness of Spreadsheet Programs' Annotation-less Unit Type Inference for C Tudor Antoniu<sup>†</sup> Paul A. Steckler<sup>‡</sup> Shriram Krishna Philip Guo and Stephen McCamant ınk Microsystems Northrop Grumman IT/FNMOC Brown Unive **Dimensions and Units** Final Project, 6.883: Program Analysis Erich Neuwirth Matthias Felleisen December 14, 2005 Automatic Dimensional Inference Rule-based Analysis of Dimensional Safety Ab Mitchell Wand\* Patrick O'Keefe scientific Feng Chen, Grigore Roşu, Ram Prasad Venkatesan companies, er y include increasingly l Department of Computer Science rams are ıms. The create ICAD, Inc. College of Computer Science University of Illinois at Urbana - Champaign, USA errors and must track then {fengchen,grosu,rpvenkat}@uiuc.edu Northeastern University 360 Huntington Avenue, 161CN Cambridge, MA 02139 Abstract. Dimension analysis concerned wi Boston, MA 02115, USA Inférence d'unités physiques en ML ciples of units of mea wand@corwin.ccs.northeastern.edu routinely dimensional can hide significant of to find otherwise. Di $Jean Goubault^{1,2}$

Bull coordination recherche

78 340 Les Clayes sous Bois, France

DMI-LIENS Ecole Normale Supéri

Edit this pad

Jean.Goubault@frcl.bull.f

rue Jean Jaurès

tional programming eral design principles prototypes, impleme static checkers. Our code which are prope programming langua types consists of war safety policy. These Maude, using more non-trivial applicatio

#### 1 Introduction

Checking software for me analysis, is an old topic in mains, such as physics, me involves units of measurem programming languages. units can be quite comple putations, for example add domain-specific errors which

Résumé : Nous décrivons une extension du systè un typage plus fin des quantités numériques, par physique (masse, longueur, etc.). Le système est effectue la vérification et l'inférence automatique des physiques (kg, m, etc.) sont alors des échelles le long de

automatiquement les instructions de conversion ent Nous en décrivons les principes, la réalisation

# idea!

Not a new

1000 Massachusetts Avenue

TypeRef DimRef TypeRef · DimRef TypeRef / DimRef TypeRef per DimRef TypeRef UnitRef TypeRef · UnitRef TypeRef / UnitRef TypeRef per UnitRef TypeRef in DimRef StaticArg Unity dimensionless

DimType

DimRef

StaticArg · StaticArg StaticArg StaticArg Arg / StaticArg

> aticArg Arg ^ StaticArg Arg per StaticArg

eOp StaticArg Arg DUPostOp

ta types for performing arithmetic with physical he physical dimensions of the quantities/units of operations is verified by the type axing of numerical values as quantities mits. The library is designed to, as far as of unit usage.

Issues

Not logged in Log in | Help

1. While there have been a number of proposals to integrate dimensional

sy observation that dimensional analysis fits neatly into the pattern

le type inference [4, 5, 6]. In this paper we show how to add

the simply-typed lambda calculus, and we show that every

mensional

tatically checked physical

mensions for Haskell.

Wiki

n-preserving term has a principal type. The principal type

analysis into existing compilers [1, 7, 8, 9], it appears that no one has made

Related changes

#### Adding Apples and Orange

Martin Erwig and Margaret Burnett

Oregon State University Department of Computer Science Corvallis, OR 97331, USA [erwig|burnett]@cs.orst.edu

Abstract. We define a unit system for end-user spreadsheets that based on the concrete notion of units instead of the abstract conceptypes. Units are derived from header information given by spreadshe The unit system contains concepts, such as dependent units, mult units, and unit generalization, that allow the classification of spre sheet contents on a more fine-grained level than types do. Also, because communication with the end user happens only in terms of objects t are contained in the spreadsheet, our system does not require end users to learn new abstract concepts of type systems.

Keywords: First-Order Functional Language, Spreadsheet, Type Checking, Unit, End-User Programming

#### Dimensionalized numbers

Categories: Mathematics | Type-level SOURCEFORGE.NET

I have created a simple toy example using functional d types to do compile-time unit analysis error catching a only two "base dimensions" time, and length, and very but it is usable.

The Units of Measure Library

Provides a C++ type-safe mechanism to deal with various units of measure. It prevents many units-related run-time errors (such as mistakenly mixing feet and meters) by catching them at compile time. The library includes scalar, 2D, and 3D vectors.

#### Programming Languages and Dimensions

Andrew John Kennedy St. Catharine's College



A dissertation submitted to the University of Cambridge towards the degree of Doctor of Philosophy November 1995



MSDN → Developer Centres → Microsoft F# Developer Center → Home



F# is a functional programming language for the .NET Fr ework. It combines the succinct, expressive, and compositional st libraries, interoperability, and object model of .NET.

#### Getting Started with F#

#### Download the F# CTP

Get the newest release of F#, including the compiler, tools, as Visual Studio 2008 integration to get started developing

#### Learn F#

Get resources for learning F#, including articles, videos, and books. Three sample chapters of the Expert F≠ book are also available for preview.

#### The F# Language Specification

Get all the nitty-gritty details of the F# language from the draft F# language specification. Provides a indepth description of the F# language's syntax and semantics. Also available in PDF.

...put into practice at last!

ncement

on Syme describes the key new

into the new world of F#.

More...

#### Featured Videos







#### Talk overview

- Practice
  - What is F#?
  - A tour of units in F#
  - Case studies
- Theory
  - Type system
  - Type inference
- Future

#### What is F#?

- It's a functional language in the ML tradition core is compatible with core of Caml
  - + .NET object model , builds on experience of SML.NET, MLj
  - + active patterns, quotations, monad comprehensions, units-of-measure, lightweight syntax and other features
- Shipping as a product with next release of Visual Studio
  - Community Tech Preview released September 08
  - Also available for Mac/Linux via the Mono runtime
  - Come to Don Syme's CUFP talk (9am Friday), or the DEFUN tutorial (Sunday pm)

#### Units-of-measure in F#

- Type system extension
  - Not just a static analysis tool
- Minimally invasive
  - Type inference, in the spirit of ML & Haskell
    - Annotate literals with units, let inference do the rest
    - But overloading must be resolved
  - No run-time cost (erasure)
- Support F# object model as far as possible
- Extensible
  - Not just for floats!

### **Feature Tour**

#### Case studies

- We've been using the units feature at Microsoft for a few months now
  - Machine learning (Ralf Herbrich)
  - Games (Phil Trelford)
  - Physics simulation (Philipp Hennig, Don Syme, Chris Smith)
  - Finance (Luca Bolognese)

#### Feedback from users

- Units are useful
  - They really do catch unit errors (Ralf, Phil, Philipp)
  - They inform the developer, and "correct" types help catch errors e.g.

```
let doublesqr x = sqr x + x
     val doublesqr: float -> float

let doublesqr x = sqr x + sqr x
     val doublesqr: float<'u> -> float<'u^2>
```

- Automatic unit conversions: would be nice, but surprisingly not a big request
- Need for "unit asserts" for external code e.g.

```
type System.Math =
  with
   val Sqrt : float<'u^2> -> float<'u>
  end
```

## Theory

## The type system, informally

- Take the ML type system with Hindley-Milner inference
- Add a new sort: Measure

```
[<Measure>] type kg
type ([<Measure>] 'a) complex = {real:float<'a>; imag:float<'a>}
```

Add operators on Measures (product, inverse, no units)

```
val (*) : float<'a> -> float<'b> -> float<'a 'b>
[<Measure>] type Hz = s^-1 inverse product
val norm : Vector<'a> -> Vector<1>
```

- Build in equational theory on Measures (commutativity, associativity, identity, inverses i.e. Abelian group)
- Refine the types of arithmetic operators e.g.

```
val sqrt : float<'a^2> -> float<'a>
val (/) : float<'u> -> float<'v> -> float<'u/'v>
```

### Toy type system, formally

#### Syntax

```
units \mu ::= u \mid b \mid \mathbf{1} \mid \mu_1 \cdot \mu_2 \mid \mu^{-1}

types \tau ::= \alpha \mid \text{float } \mu \mid \tau_1 \to \tau_2 \mid \text{bool } \mid \cdots

type schemes \sigma ::= \forall \vec{u}.\tau

expressions e ::= x \mid c \in \mathbb{R} \mid e \mid a \mid \lambda x.e \mid \text{let } x = e_1 \text{ in } e_2 \mid \cdots

judgments u_1, \dots, u_m; x_1 : \sigma_1, \dots, x_n : \sigma_n \vdash e : \tau
```

#### Equational theory of units (Abelian groups)

identity 
$$\mathbf{1} \cdot \mu =_U \mu$$
 inverse  $\mu \cdot \mu^{-1} =_U \mathbf{1}$  assoc  $(\mu_1 \cdot \mu_2) \cdot \mu_3 =_U \mu_1 \cdot (\mu_2 \cdot \mu_3)$  comm  $\mu_1 \cdot \mu_2 =_U \mu_2 \cdot \mu_1$ 

#### Typing rules (excerpt)

$$\frac{(x:\forall \vec{u}.\tau) \in \Gamma}{\Delta; \Gamma \vdash x: \tau[\vec{\mu}/\vec{u}]} \qquad \frac{\Delta; \Gamma \vdash e: \tau_1}{\Delta; \Gamma \vdash e: \tau_2} \qquad \frac{\tau_1 =_U \tau_2}{\Delta; \Gamma \vdash c: \mathtt{float} \ \mathbf{1}}$$

$$\frac{\Delta, \vec{u}; \Gamma \vdash e_1: \tau_1}{\Delta; \Gamma \vdash \mathtt{let} \ x = e_1 \ \mathtt{in} \ e_2: \tau_2} \qquad \frac{\Delta; \Gamma \vdash c: \mathtt{float} \ \mathbf{1}}{\Delta; \Gamma \vdash \lambda x.e: \tau_1 \to \tau_2}$$

## Type inference and principal types

- The type systems of SML, Caml, Haskell and F# have (in principle, at least) the principal types property:
  - if expression e is typeable there exists a unique type scheme  $\sigma$  such that all valid types are instances of  $\sigma$
  - moreover, an inference algorithm will find the principal type
- If type checking e produces a type scheme that instantiates to  $\tau$  write  $tc(e) \leq \tau$
- We can express correctness as
  - − Soundness:  $tc(e) \le \tau \Rightarrow \vdash e : \tau$
  - − Completeness:  $\vdash e : \tau \Rightarrow \mathsf{tc}(e) \leq \tau$
- Units-of-measure also have principal types, but algorithm is trickier

### ML type inference algorithm

- Two essential ingredients
  - **1.** Unification. A *unifier* of two types  $\tau_1$  and  $\tau_2$  is a substitution S on type variables such that  $S(\tau_1)=S(\tau_2)$ . For unifiable types, there is a *most general unifier*.
  - **2. Generalization**. To type let  $x = e_1$  in  $e_2$  find a type  $\tau$  for  $e_1$  and then quantify on the variables that are free in  $\tau$  but not free in the type environment  $\Gamma$ .

### The good news...

- For units, a unifier of two unit expressions  $\mu_1$  and  $\mu_2$  is a substitution S on unit variables such that  $S(\mu_1) = U S(\mu_2)$
- Fortunately, Abelian Group unification is
  - unitary (unique most general unifiers exist with respect to the equational theory), and
  - decidable (algorithm is a variation of Gaussian elimination)

### Unification algorithm

```
Unify(\mu_1, \mu_2) = \text{UnifyOne}(\mu_1 \cdot \mu_2^{-1})
UnifyOne(\mu) =
   let \mu = u_1^{x_1} \cdots u_m^{x_m} \cdot b_1^{y_1} \cdots b_n^{y_n} where |x_1| \leq |x_2|, \cdots, |x_m|
   in
       if m = 0 and n = 0 then I
       if m = 0 and n \neq 0 then fail
       if m = 1 and x_1 | y_i for all i then \{u_1 \mapsto b_1^{-y_1/x_1} \cdots b_m^{-y_n/x_1}\}
       if m=1 otherwise then fail
       else S_2 \circ S_1 where
           S_1 = \{u_1 \mapsto u_1 \cdot u_2^{-\lfloor x_2/x_1 \rfloor} \cdots u_m^{-\lfloor x_m/x_1 \rfloor} \cdot b_1^{-\lfloor y_1/x_1 \rfloor} \cdots b_n^{-\lfloor y_n/x_1 \rfloor} \}
           S_2 = \text{UnifyOne}(S_1(\mu))
```

#### Unification in action

$$u^3 \cdot v^2 =_U \mathrm{kg}^6$$
 $\mathbf{u}^3 \cdot v^2 \cdot \mathrm{kg}^{-6} =_U \mathbf{1}$ 
 $\mathbf{u}$  apply  $\{v \mapsto v \cdot u^{-1} \cdot \mathrm{kg}^3\}$ 
 $u \cdot v^2 =_U \mathbf{1}$ 
 $\mathbf{u}$  apply  $\{u \mapsto u \cdot v^{-2}\}$ 
 $u =_U \mathbf{1}$ 
 $\mathbf{u}$  apply  $\{u \mapsto \mathbf{1}\}$ 
 $\mathbf{1} =_U \mathbf{1}$ 
Success!

#### The bad news...

- Generalization based on free variables is sound but not complete for units-of-measure
- Why? Because the notion of syntactic free variables is not stable under various transformations.
  - 1. "Free variables" is not stable under equivalence of types e.g.  $fv(u \cdot v \cdot u^{-1}) \neq fv(v)$ . Solution: normalize
  - 2. "Free variables" is not stable under equivalence of type schemes e.g. fv (∀ u. float<u· v>→ float<u· v>) ≠ fv(∀ u.float<u>→ float<u>) Solution: normalize
  - 3. "Generalizable variables" is not stable under equivalence of typings e.g.

```
d: float< u \cdot v > \vdash expr: float< u \cdot w > \rightarrow float< v \cdot w^{-1} > Solution: normalize
```

## Type Scheme Equivalence

- Two type schemes are equivalent if they instantiate to the same set of types
- For vanilla ML, this just amounts to renaming quantified type variables or removing redundant quantifiers.
- For ML + units, there are many non-trivial equivalences. E.g.

```
\begin{array}{l} /: \forall uv. \texttt{float} \ u \to \texttt{float} \ v \to \texttt{float} \ u \cdot v^{-1} \\ /: \forall uvw. \texttt{float} \ w \cdot u \to \texttt{float} \ v \to \texttt{float} \ w \cdot u \cdot v^{-1} \\ /: \forall uv. \texttt{float} \ u^{-1} \to \texttt{float} \ v^{-1} \to \texttt{float} \ u^{-1} \cdot v \\ /: \forall uv. \texttt{float} \ u \cdot v \to \texttt{float} \ u \to \texttt{float} \ v \\ /: \forall uv. \texttt{float} \ u \to \texttt{float} \ v \to \texttt{float} \ u \cdot v \\ /: \forall uv. \texttt{float} \ w \cdot u \to \texttt{float} \ w \cdot v \to \texttt{float} \ u \cdot v^{-1} \end{array}
```

## Simplifying type schemes

- It's possible to show that two type schemes are equivalent iff there
  is an invertible substitution on the bound variables that maps
  between them (this is a "change of basis")
- *Idea*: compute such a substitution that puts a type scheme in some kind of preferred "normal form". Desirable properties:
  - No redundant bound or free variables (so number of variables = number of "degrees of freedom")
  - Minimize size of exponents
  - Use positive exponents if possible
  - Unique up to renaming
- Such a form does exist, and corresponds to Hermite Normal Form from algebra
  - Pleasant side-effect: deterministic ordering on variables in type

## Simplification in action

## Generalization: example problem

Suppose

$$\Gamma = \{/ : \forall uv. \mathtt{float} \ u \cdot v \to \mathtt{float} \ u \to \mathtt{float} \ v \}$$

Infer a type for

$$\emptyset$$
;  $\Gamma \vdash \lambda x$ .let  $f = /x$  in  $(f \ 1.0 < kg >, f \ 2.0 < s >) : ?$ 

Problem: when typing let, we can't generalize:

$$\Gamma, x : \mathtt{float}\, u \cdot v \vdash /x : \mathtt{float}\, u \to \mathtt{float}\, v$$

Solution: apply a "simplifying" substitution to the environment:

$$u\mapsto u\cdot v^{-1}$$

$$\Gamma, x : \mathtt{float}\, u \vdash /x : \mathtt{float}\, u \cdot v^{-1} \to \mathtt{float}\, v$$

#### Generalization

#### Recipe:

- Use the "simplify" algorithm on the *free variables* of the type environment  $\Gamma$  to compute an invertible change-of-basis substitution S
- Apply S to both \( \Gamma\) and the inferred type \( \tau\)
- Compute generalizable variables in the usual way i.e.  $fv(S(\tau)) \setminus fv(S(\Gamma))$
- Apply S<sup>-1</sup> to the resulting type scheme.

#### Summary:

Gen
$$(\Gamma, \tau) = S^{-1}(\forall u_1, \dots, u_n. S(\tau))$$
  
where  $\text{fv}(S(\tau)) \setminus \text{fv}(S(\Gamma)) = \{u_1, \dots, u_n\}$   
and  $S$  is simplifier of free variables of  $\Gamma$ .

#### Generalization in action

$$\Gamma, x: \texttt{float} \ u \cdot v \vdash /x: \texttt{float} \ u \to \texttt{float} \ v$$
 
$$\downarrow u \mapsto u \cdot v^{-1}$$
 
$$\Gamma, x: \texttt{float} \ u \vdash /x: \texttt{float} \ u \cdot v^{-1} \to \texttt{float} \ v$$
 
$$\downarrow \texttt{quantify}$$
 
$$\Gamma, x: \texttt{float} \ u \vdash /x: \forall v. \texttt{float} \ u \cdot v^{-1} \to \texttt{float} \ v$$
 
$$\downarrow \texttt{rename}$$
 
$$\Gamma, x: \texttt{float} \ u \vdash /x: \forall w. \texttt{float} \ u \cdot w^{-1} \to \texttt{float} \ w$$
 
$$\downarrow u \mapsto u \cdot v$$
 
$$\Gamma, x: \texttt{float} \ u \cdot v \vdash /x: \forall w. \texttt{float} \ u \cdot v \cdot w^{-1} \to \texttt{float} \ w$$

## Beyond Hindley-Milner

#### Non-regular datatypes

```
// Non-regular datatype: a list of derivatives of a function
type derivs<[<Measure>] 'u, [<Measure>] 'v> =
  Nil
| Cons of (float<'u> -> float<'v>) * derivs<'u,'v/'u>
```

#### Polymorphic recursion

```
let rec makeDerivs<[<Measure>] 'u, [<Measure>] 'v>
  (h:float<'u>)
  (f:float<'u> -> float<'v>)
  (n:int) : derivs<'u,'v> =
  if n=0 then Nil else Cons(f, makeDerivs h (diff h f) (n-1))
```

## Future developments (F# v.next?)

- Automatic unit conversion
  - FAQ, but not a top priority for developers who have tried out the feature
  - Implicit insertion of floating-point operations considered harmful?
- Units for external code: asserting a type
  - Should be controlled: not a general cast mechanism
- Higher kinds e.g. Consider

```
type Matrix<[<Measure>] 'u, 't> = 't<'u> array array
```

#### Conclusion

- Units-of-measure types occupy a "sweet spot" in the space of type systems
  - Type system is easy to understand for novices
  - Typing rules are very simple
  - Types have a simple form (e.g. no constrained polymorphism)
  - Types don't intrude (there is rarely any need for annotation)
  - Behind the scenes, inference is non-trivial but practical

#### **Pointers**

- F# download: http://msdn.microsoft.com/fsharp
- Blog on units: <a href="http://blogs.msdn.com/andrewkennedy">http://blogs.msdn.com/andrewkennedy</a>
- Thesis and papers: http://research.microsoft.com/~akenn/units